

BEST AVAILABLE COPY

FIG. 1A

SEQ ID NO: 1

pcmv II

→ 1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG  
AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

---

51 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCC  
CTCTGCCAGT GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTGCGGGC

---

101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG  
AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

---

HindIII  
-----

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGTCA  
GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT

---

201 AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTAATTCTGG AATAGCTCAG  
TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

---

251 AGGCGGAGGC GGCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA  
TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

---

301 TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG  
ACCCCGCCTC TTACCCGCCT TGACCCGCCC CTCCCTTAAT AACCGATAAC

---

351 GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

---

401 CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

---

451 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTCCG  
ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

---

501 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC  
GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

---

551 CCCGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA  
GGGCGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

---

601 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAAGTGGCCA  
CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

---

651 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG  
GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

---

701 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA  
AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GACTGGAAT

---

751 CGGGACTTTC TACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

---

801 CATGGTGATG CGGTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG  
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC

---

851 ACTCACGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG  
TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC

FIG. 1B

901	TTTTGGCACC	AAAATCAACG	GGACTTTCCA	AAATGTCGTA	ATAACCCCGC
	AAAACCGTGG	TTTtagTtGC	CCTGAAAGGT	TTTACAGCAT	TATTGGGGCG
951	CCCGTTGACG	CAAATGGGCG	GtagGCGTGT	ACGGTGGGAG	GTCTATATAA
	GGGCAACTGC	GTTTACC CGC	CATCCGCACA	TGCCACCCTC	CAGATATATT
1001	GCAGAGCTCG	TTTAgTGAAC	CGTCAGATCG	CCTGGAGACG	CCATCCACGC
	CGTCTCGAGC	AAATCACTTG	GCAGTCTAGC	GGACCTCTGC	GGTAGGTGCG
1051	TGTTTTGACC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGGCCG
	ACAAAAC TGG	AGGTATCTTC	TGTGGCCCTG	GCTAGGTCCG	AGGCGCCGGC
1101	GGAACGGTGC	ATTGGAACGC	GGATTCCCCG	TGCCAAGAGT	GACGTAAGTA
	CCTTGCCACG	TAACCTTGCG	CCTAAGGGGC	ACGGTTCTCA	CTGCATT CAT
1151	CCGCCTATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA
	GGCGGATATC	TGAGATATCC	GTGTGGGGAA	ACCGAGAATA	CGTACGATAT
1201	CTGTTTTTGG	CTTGGGGCCT	ATACACCCCC	GCTCCTTATG	CTATAGGTGA
	GACAAAAACC	GAACCCCGGA	TATGTGGGGG	CGAGGAATAC	GATATCCACT
1251	TGGTATAGCT	TAGCCTATAG	GTGTGGGTTA	TTGACCATTA	TTGACCACTC
	ACCATATCGA	ATCGGATATC	CACACCCAAT	AACTGGTAAT	AACTGGTGAG
1301	CCCTATTGGT	GACGATACTT	TCCATTACTA	ATCCATAACA	TGGCTCTTTG
	GGGATAACCA	CTGCTATGAA	AGGTAATGAT	TAGGTATTGT	ACCGAGAAAC
1351	CCACAAC TAT	CTCTATTGGC	TATATGCCAA	TACTCTGTCC	TTCAGAGACT
	GGTGTGATA	GAGATAACCG	ATATACGGTT	ATGAGACAGG	AAGTCTCTGA
1401	GACACGGACT	CTGTATTTTT	ACAGGATGGG	GTCCATT TAT	TATTTACAAA
	CTGTGCCTGA	GACATAAAAA	TGTCTTACCC	CAGGTAAATA	ATAAATGTTT
1451	TTACATATA	CAACAACGCC	GTCCCCCGTG	CCCGCAGTTT	TTATTAAACA
	AAGTGTATAT	GTTGTTGCGG	CAGGGGGCAC	GGGCGTCAAA	AATAATT TGT
1501	TAGCGTGGGA	TCTCCGACAT	CTCGGGTACG	TGTTCCGGAC	ATGGGCTCTT
	ATCGCACCTT	AGAGGCTGTA	GAGCCCATGC	ACAAGGCC TG	TACCCGAGAA
1551	CTCCGGTAGC	GGCGGAGCTT	CCACATCCGA	GCCCTGGTCC	CATCCGTCCA
	GAGGCCATCG	CCGCCTCGAA	GGTGTAGGCT	CGGGACCAGG	GTAGGCAGGT
1601	GCGGCTCATG	GTCGCTCGGC	AGCTCCTTGC	TCCTAACAGT	GGAGGCCAGA
	CGCCGAGTAC	CAGCGAGCCG	TCGAGGAACG	AGGATTGTCA	CCTCCGGTCT
1651	CTTAGGCACA	GCACAATGCC	CACCACCACC	AGTGTGCCGC	ACAAGGCCGT
	GAATCCGTGT	CGTGTTACGG	GTGGTGGTGG	TCACACGGCG	TGTTCCGGCA
1701	GGCGGTAGGG	TATGTGTCTG	AAAATGAGCT	CGGAGATTGG	GCTCGCACCT
	CCGCCATCCC	ATACACAGAC	TTTTACTCGA	GCCTCTAACC	CGAGCGTGGA
1751	GGACGCAGAT	GGAAGACTTA	AGGCAGCGGC	AGAAGAAGAT	GCAGGCAGCT
	CCTGCGTCTA	CCTCTGGAAT	TCCGTCGCCG	TCTTCTTCTA	CGTCCGTCGA
1801	GAGTTGTTGT	ATTCTGATAA	GAGTCAGAGG	TAACTCCCGT	TGCGGTGCTG
	CTCAACAACA	TAAGACTATT	CTCAGTCTCC	ATTGAGGGCA	ACGCCACGAC

FIG. 1C

1851	TTAACGGTGG	AGGGCAGTGT	AGTCTGAGCA	GTACTCGTTG	CTGCCGCGCG
	AATTGCCACC	TCCCGTCACA	TCAGACTCGT	CATGAGCAAC	GACGGCGCGC
<hr/>					
1901	CGCCACCAGA	CATAATAGCT	GACAGACTAA	CAGACTGTTT	CTTTCCATGG
	GCGGTGGTCT	GTATTATCGA	CTGTCTGATT	GTCTGACAAG	GAAAGGTACC
<hr/>					
		SalI	EcoRI	XhoI	
		-----	-----	-----	
1951	GTCTTTTCTG	CAGTCACCGT	CGTCGACCTA	AGAATTCAGA	CTCGAGCAAG
	CAGAAAAGAC	GTCACTGGCA	GCAGCTGGAT	TCTTAAGTCT	GAGCTCGTTC
<hr/>					
	XbaI	AscI	EcoRV	BamHI	MluI
	-----	-----	-----	-----	-----
2001	TCTAGAAAGG	CGCGCCAAGA	TATCAAGGAT	CCACTACGCG	TTAGAGCTCG
	AGATCTTTCC	GCGCGTTTCT	ATAGTTCCTA	GGTGATGCGC	AATCTCGAGC
<hr/>					
2051	CTGATCAGCC	TCGACTGTGC	CTTCTAGTTG	CCAGCCATCT	GTTGTTTGCC
	GACTAGTCGG	AGCTGACACG	GAAGATCAAC	GGTCGGTAGA	CAACAAACGG
<hr/>					
2101	CCTCCCCCGT	GCCTTCCTTG	ACCCTGGAAG	GTGCCACTCC	CACTGTCCTT
	GGAGGGGGCA	CGGAAGGAAC	TGGGACCTTC	CACGGTGAGG	GTGACAGGAA
<hr/>					
2151	TCCTAATAAA	ATGAGGAAAT	TGCATCGCAT	TGTCTGAGTA	GGTGTCATTC
	AGGATTATTT	TACTCCTTTA	ACGTAGCGTA	ACAGACTCAT	CCACAGTAAG
<hr/>					
2201	TATTCTGGGG	GGTGGGGTGG	GGCAGGACAG	CAAGGGGGAG	GATTGGGAAG
	ATAAGACCCC	CCACCCACCC	CCGTCTGTGC	GTTCCCCCTC	CTAACCCCTC
<hr/>					
2251	ACAATAGCAG	GCATGCTGGG	GAGCTCTTCC	GCTTCCTCGC	TAAGTACTGC
	TGTTATCGTC	CGTACGACCC	CTCGAGAAGG	CGAAGGAGCG	AGTGACTGAG
<hr/>					
2301	GCTGEGCTCG	GTGTTTCGGC	TGCGGCGAGC	GGTATCAGCT	CACTCAAAGG
	CGACGCGAGC	CAGCAAGCCG	ACGCCGCTCG	CCATAGTCGA	GTGAGTTTCC
<hr/>					
2351	CGGTAATACG	GTTATCCACA	GAATCAGGGG	ATAACGCAGG	AAAGAACATG
	GCCATTATGC	CAATAGGTGT	CTTAGTCCCC	TATTGCGTCC	TTTCTTGATC
<hr/>					
2401	TGAGCAAAAG	GCCAGCAAAA	GGCCAGGAAC	CGTAAAAAGG	CCGCGTTGCT
	ACTCGTTTTT	CGGTCTTTTT	CCGGTCTTTG	GCATTTTTCC	GGCGCAACGA
<hr/>					
2451	GGCGTTTTTC	CATAGGCTCC	GGCCCCCTGA	CGAGCATCAC	AAAAATCGAC
	CCGCAAAAAG	GTATCCGAGG	CGGGGGGACT	GCTCGTAGTG	TTTTTAGCTG
<hr/>					
2501	GCTCAAGTCA	GAGGTGGCGA	AACCCGACAG	GACTATAAAG	ATACCAGGCG
	CGAGTTCACT	CTCCACCGCT	TTGGGCTGTC	CTGATATTTT	TATGGTCCGC
<hr/>					
2551	TTTCCCCCTG	GAAGCTCCCT	CGTGCGCTCT	CCTGTTCCGA	CCCTGCCGCT
	AAAGGGGGAC	CTTCGAGGGA	GCACGCGAGA	GGACAAGGCT	GGGACGGCGA
<hr/>					
2601	TACCGGATAC	CTGTCCGCTT	TTCTCCCTTC	GGGAAGCGTG	GCGCTTTCTC
	ATGGCCTATG	GACAGGCGGA	AAGAGGGAAG	CCCTTCGCAC	CGCGAAAGAG
<hr/>					
2651	AATGCTCACG	CTGTAGGTAT	CTCAGTTCGG	TGTAGGTCTG	TCGCTCCAAG
	TTACGAGTGC	GACATCCATA	GAGTCAAGCC	ACATCCAGCA	AGCGAGGTTC
<hr/>					
2701	CTGGGCTGTG	TGCACGAACC	CCCCGTTCAG	CCCGACCGCT	GCGCCTTATC
	GACCCGACAC	ACGTGCTTGG	GGGGCAAGTC	GGGCTGGCGA	CGCGGAATAG

FIG. 1D

2751 CCGTAACTAT CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATCGCCAC  
GCCATTGATA GCAGAACTCA GGTGGGCCA TTCTGTGCTG AATAGCGGTG

2801 TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT  
ACCGTCGTCG GTGACCATTG TCCTAATCGT CTCGCTCCAT ACATCCGCCA

2851 GCTACAGAGT TCTTGAAGTG GTGGCCTAAC TACGGCTACA CTAGAAGGAC  
CGATGTCTCA AGAACTTCAC CACCGGATTG ATGCCGATGT GATCTTCCTG

2901 AGTATTTGGT ATCTGCGCTC TGCTGAAGCC AGTTACCTTC GGAAAAAGAG  
TCATAAACCA TAGACGCGAG ACGACTTCGG TCAATGGAAG CCTTTTCTC

2951 TTGGTAGCTC TTGATCCGGC AAACAAACCA CCGCTGGTAG CCGTGGTTTT  
AACCATCGAG AACTAGGCCG TTTGTTTGGT GCGGACCATC GCCACCAAAA

3001 TTTGTTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAAGAAGA  
AAACAAACGT TCGTCGTCTA ATGCGCGTCT TTTTTCCTA GAGTCTCTCT

3051 TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC  
AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTTGAGTG

3101 GTTAAGGGAT TTTGGTCATG AGATTATCAA AAAGGATCTT CACCTAGATC  
CAATTCCTA AAACCAGTAC TCTAATAGTT TTTCTAGAA GTGGATCTAG

3151 CTTTTAAATT AAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA  
GAAAATTTAA TTTTACTTC AAAATTTAGT TAGATTTTAT ATATACTCAT

3201 AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG  
TTGAACCAGA CTGTCAATGG TTACGAATTA GTCACTCCGT GGATAGAGTC

3251 CGATCTGTCT ATTTGTTTCA TCCATAGTTG CCTGACTCCC CGTCGTGTAG  
GCTAGACAGA TAAAGCAAGT AGGTATCAAC GGACTGAGGG GCAGCACATC

3301 ATAACCTACGA TACGGGAGGG CTTACCATCT GGCCCCAGTG CTGCAATGAT  
TATTGATGCT ATGCCCTCCC GAATGGTAGA CCGGGGTCAC GACGTTACTA

3351 ACCGCGAGAC CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC  
TGGCGCTCTG GGTGCGAGTG GCCGAGGTCT AAATAGTCGT TATTTGGTCG

3401 CAGCCGGAAG GGCCGAGCGC AGAAGTGGTC CTGCAACTTT ATCCGCCCTCC  
GTCCGCCCTC CCGGCTCGCG TCTTCACCAG GACGTTGAAA TAGGCGGAGG

3451 ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT  
TAGGTCAGAT AATTAACAAC GGCCCTTCGA TCTCATTAT CAAGCGGTCA

3501 TAATAGTTTG CGCAACGTTG TTGCCATTGC TACAGGCATC GTGGTGTAC  
ATTATCAAAC GCGTTGCAAC AACGTAACG ATGTCCGTAG CACCACAGTG

3551 GCTCGTCGTT TGGTATGGCT TCATTAGCT CCGGTTCCCA ACGATCAAGG  
CGAGCAGCAA ACCATACCGA AGTAAGTCGA GGCCAAGGGT TGCTAGTTCC

3601 CGAGTTACAT GATCCCCCAT GTTGTGCAAA AAAGCGGTTA GCTCCTTCGG  
GCTCAATGTA CTAGGGGGTA CAACACGTTT TTTGCGCAAT CGAGGAAGCC

3651 TCCTCCGATC GTTGTGAGAA GTAAGTTGGC CGCAGTGTTA TCACTCATGG  
AGGAGGCTAG CAACAGTCTT CATCAACCG GCGTCACAAT AGTGAGTACC

FIG. 1E

3701 TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC  
 AATACCGTCG TGACGTATTA AGAGAATGAC AGTACGGTAG GCATTCTACG

---

3751 TTTTCTGTGA CTGGTGAGTA CTCACCAAG TCATTCTGAG AATAGTGTAT  
 AAAAGACACT GACCACTCAT GAGTTGGTTC AGTAAGACTC TTATCACATA

---

3801 GCGGCGACCG AGTTGCTCTT GCCCGGCGTC AATACGGGAT AATACCGCGC  
 CGCCGCTGGC TCAACGAGAA CGGGCCGCAG TTATGCCCTA TTATGGCGCG

---

3851 CACATAGCAG AACTTTAAAA GTGCTCATCA TTGGAAAACG TTCTTCGGGG  
 GTGTATCGTC TTGAAATTTT CACGAGTAGT AACCTTTTGC AAGAAGCCCC

---

3901 CGAAAACTCT CAAGGATCTT ACCGCTGTTG AGATCCAGTT CGATGTAACC  
 GCTTTTGAGA GTTCCTAGAA TGGCGACAAC TCTAGGTCAA GCTACATTGG

---

3951 CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT  
 GTGAGCACGT GGGTTGACTA GAAGTCGTAG AAAATGAAAG TGGTCGCAAA

---

4001 CTGGGTGAGC AAAAACAGGA AGGCAAAATG CCGCAAAAAA GGGAATAAGG  
 GACCCACTCG TTTTGTCCCT TCCGTTTAC GCGTTTTTT CCCTTATTCC

---

4051 GCGACACGGA AATGTTGAAT ACTCATACTC TTCCTTTTTC AATATTATTG  
 CGCTGTGCCT TTACAACTTA TGAGTATGAG AAGGAAAAAG TTATAATAAC

---

4101 AAGCATTTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA  
 TTCGTAAATA GTCCCAATAA CAGAGTACTC GCCTATGTAT AAACCTACAT

---

4151 TTTAGAAAAA TAAACAAATA GGGGTTCGCG GCACATTTC CCGAAAAGTG  
 AAATCTTTTT ATTTGTTTAT CCCCAAGGCG CGTGTAAGG GGCTTTTCAC

---

4201 CCACCTGACG TCTAAGAAAC CATTATTATC ATGACATTAA CCTATAAAAA  
 GGTGGACTGC AGATTCTTTG GTAATAATAG TACTGTAATT GGATATTTTT

---

4251 TAGGCGTATC ACGAGGCCCT TTCGTC  
 ATCCGCATAG TGCTCCGGGA AAGCAG

---

FIG. 1F

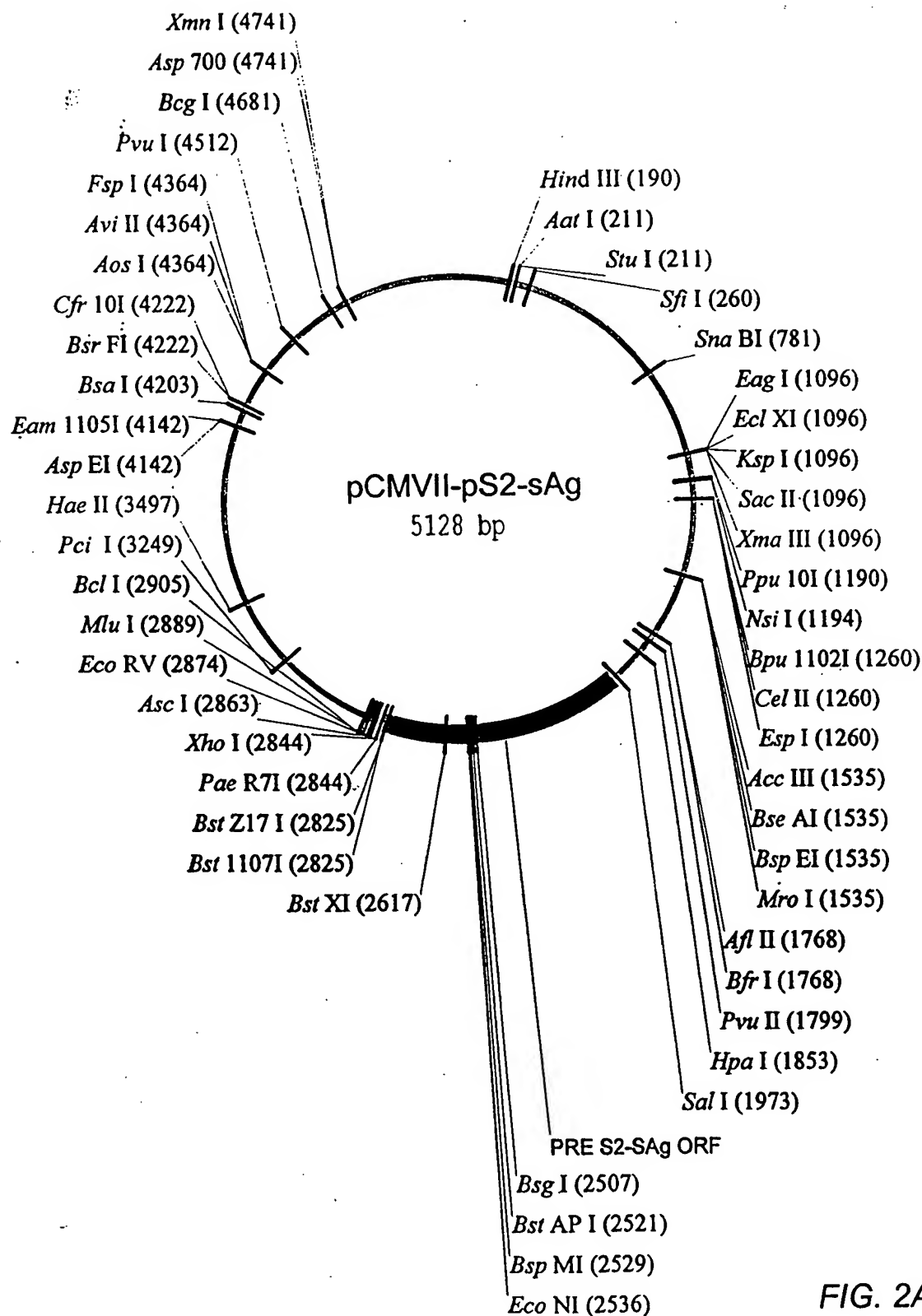


FIG. 2A

SEQ ID NO: 271

TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG  
AGCGCGCAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

51 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG  
CTCTGCCAGT GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTGCGGC

101 TCAGGGCGGG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG  
AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

HindIII

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA  
GCCGTAGTCT CGTCTAACAT GACTCTCAG TGGTATACTT CGAAAAACGT

StuI

AatI

201 AAAGCCTAGG CCTCCAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG  
TTTCGGATCC GGAGGTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

SfiI

251 AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA  
TCCGGCTCCG CCGGAGCCCG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

301 TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG  
ACCCCGCCTC TTACCCGCCT TGACCCGCCC CTCCCTTAAT AACCGATAAC

351 GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401 CATGTCCAAT ATGACCGCCA TGTGACATT GATTATTGAC TAGTTATTAA  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

451 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTCCG  
ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

501 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC  
GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

551 CCGGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA  
GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

601 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAAGTGGCCA  
CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

651 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG  
GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

701 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA  
AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

SnaBI

751 CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

FIG. 2B



801 CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG  
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC

851 ACTCAGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG  
TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC

901 TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC  
AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG

951 CCCGTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA  
GGGCAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT

1001 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC  
CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG

XmaIII

SacII

KspI

EclXI

EagI

1051 TGTTTGGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG  
ACAAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC

1101 GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA  
CCTTGCCACG TAACCTTGC CTAAGGGGC ACGGTTCTCA CTGCATTCAT

Ppu10I

NsiI

1151 CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTGGGGCCT ATACACCCCC GTCCTTATG CTATAGGTGA  
GACAAAACG GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT

EspI

CelII

Bpu1102I

1251 TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC  
ACCATATCGA ATCGGATATC CACACCCAAT AACTGGTAAT AACTGGTGAG

1301 CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG  
GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC

1351 CCACAATAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT  
GGTGTTGATA GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA

1401 GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT TATTACAAA  
CTGTGCCTGA GACATAAAAA TGCCTACCC CAGGTAAATA ATAAATGTTT

FIG. 2C

1451 TTCACATATA CAACAACGCC GTCCCCGTG CCCGAGTTT TTATTAAACA  
AAGTGTATAT GTTGTTCGGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT

MroI  
~~~~~  
BspEI  
~~~~~  
BseAI  
~~~~~  
AccIII  
~~~~~

1501 TAGCGTGGGA TCTCCGACAT CTCGGGTACG TGTTCCGGAC ATGGGCTCTT  
ATCGCACCT AGAGGCTGTA GAGCCCATGC ACAAGGCTG TACCCGAGAA

1551 CTCCGGTAGC GCGGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGGC AGTCCTTGC TCCTAACAGT GGAGGCCAGA  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

1651 CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCG ACAAGGCCGT  
GAATCCGTGT CGTGTACCG GTGGTGGTGG TCACACGGCG TGTCCGGCA

1701 GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT  
CGCCATCCC ATACACAGAC TTTACTCGA GCCTCTAACC CGAGCGTGA

BfrI  
~~~~~  
AflII  
~~~~~

PvuII  
~~~~~

1751 GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT  
CCTGCGTCTA CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA

PvuII  
~

HpaI  
~

1801 GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TCGGGTGCTG  
CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC

HpaI  
~~~~~

1851 TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG  
AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

1901 CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG  
GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

+2

SEQ ID NO: 3 → M Q W N

Sall  
~~~~~

1951 GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCATG CAGTGGAAC  
CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTAC GTCACCTTGA

+2 S T A F H Q T L Q D P R V R G L Y

2001 CCACTGCCTT CCACCAAAC CTGCAGGATC CCAGAGTCAG GGTCTGTAT  
GGTGACGGAA GGTGGTTTGA GACGTCCTAG GGTCTCAGTC CCCAGACATA

+2 L P A G G S S S G T V N P A P N I

2051 CTTCTGCTG GTGGCTCCAG TTCAGGAACA GTAAACCTG CTCCGAATAT  
GAAGGACGAC CACCGAGGTC AAGTCCTTGT CATTGGGAC GAGGCTTATA

FIG. 2D

+2 A S H I S S I S A R T G D P V T  
 2101 TGCCTCTCAC ATCTCGTCAA TCTCCGCGAG GACTGGGGAC CCTGTGACGA  
 ACGGAGAGTG TAGAGCAGTT AGAGGCGCTC CTGACCCCTG GGACACTGCT

+2 N M E N I T S G F L G P L L V L Q  
 2151 ACATGGAGAA CATCACATCA GGATTCCTAG GACCCCTGCT CGTGTACAG  
 TGTACCTCTT GTAGTGTAGT CTAAGGATC CTGGGGACGA GCACAATGTC

+2 A G F F L L T R I L T I P Q S L D  
 2201 GCGGGGTTTT TCTTGTGAC AAGAATCCTC ACAATACCGC AGAGTCTAGA  
 CGCCCCAAA AGAACAACCTG TTCTTAGGAG TGTATGGCG TCTCAGATCT

+2 S W W T S L N F L G G S P V C L  
 2251 CTCGTGGTGG ACTTCTCTCA ATTTTCTAGG GGGATCTCCC GTGTGTCTTG  
 GAGCACCACC TGAAGAGAGT TAAAGATCC CCCTAGAGGG CACACAGAAC

+2 G Q N S Q S P T S N H S P T S C P  
 2301 GCCAAAATTC GCAGTCCCCA ACCTCCAATC ACTCACCAC CTCCTGTCCT  
 CGGTTTTAAG CGTCAGGGGT TGGAGGTTAG TGAGTGGTTG GAGGACAGGA

+2 P I C P G Y R W M C L R R F I I F  
 2351 CCAATTTGTC CTGGTTATCG CTGGATGTGT CTGCGCGGTT TTATCATATT  
 GGTTAAACAG GACCAATAGC GACCTACACA GACGCCGCAA AATAGTATAA

+2 L F I L L L C L I F L L V L L D  
 2401 CCTCTTCATC CTGCTGCTAT GCCTCATCTT CTTATTGGTT CTTCTGGATT  
 GGAGAAGTAG GACGACGATA CGGAGTAGAA GAATAACCA GAAGACCTAA

+2 Y Q G M L P V C P L I P G S T T T  
 2451 ATCAAGGTAT GTTGCCCGTT TGTCCTCTAA TTCCAGGATC AACCAACAAC  
 TAGTTCCATA CAACGGGCAA ACAGGAGATT AAGGTCCTAG TTGTTGTTGG

+2 S T G P C K T C T T P A Q G N S M  
 BstAP I

~~~~~  
 BspMI

~~~~~  
 EcoNI

~~~~~  
 2501 AGTACGGGAC CATGCAAAAC CTGCACGACT CCTGCTCAAG GCAACTCTAT  
 TCATGCCCTG GTACGTTTTG GACGTGCTGA GGACGAGTTC CGTTGAGATA  
 BsgI  
 ~~~~~

+2 F P S C C C T K P T D G N C T C  
 2551 GTTTCCTCA TGTGCTGTA CAAAACCTAC GGATGGAAAT TGCACCTGTA  
 CAAAGGGAGT ACAACGACAT GTTTTGGATG CCTACCTTA ACGTGGACAT

+2 I P I P S S W A F A K Y L W E W A  
 BstXI

~~~~~  
 2601 TTCCCATCCC ATCGTCCTGG GCTTTCGCAA AATACCTATG GGAGTGGGCC  
 AAGGGTAGGG TAGCAGGACC CGAAAGCGTT TTATGGATAC CCTCACCCGG

+2 S V R F S W L S L L V P F V Q W F  
 2651 TCAGTCCGTT TCTCTTGGCT CAGTTTACTA GTGCCATTG TTCAGTGGTT  
 AGTCAGGCAA AGAGAACCGA GTCAAATGAT CACGGTAAAC AAGTCACCAA

+2 V G L S P T V W L S A I W M M W  
 2701 CGTAGGGCTT TCCCCACTG TTTGGCTTTC AGCTATATGG ATGATGTGGT  
 GCATCCCGAA AGGGGGTGAC AAACCGAAAG TCGATATACC TACTACACCA

FIG. 2E

+2 Y W G P S L Y S I V S P F I P L L  
 2751 ATTGGGGGCC AAGTCTGTAC AGCATCGTGA GTCCCTTTAT ACCGCTGTTA  
 TAACCCCGG TTCAGACATG TCGTAGCACT CAGGGAAATA TGGCGACAAT

+2 P I F F C L W V Y I \*

Bst217 I

XhoI

Bst1107I

PaeR7I

2801 CCAATTTTCT TTTGTCTCTG GGTATACATT TAAGAATTCA GACTCGAGCA  
 GGTAAAAAGA AAACAGAGAC CCATATGTAA ATTCTTAAGT CTGAGCTCGT

AscI

EcoRV

MluI

2851 AGCTAGAAA GCGCGCCAA GATATCAAGG ATCCACTACG CGTTAGAGCT  
 TCAGATCTTT CCGCGCGGTT CTATAGTTCC TAGGTGATGC GCAATCTCGA

BclI

2901 CGCTGATCAG CCTCGACTGT GCCTTCTAGT TGCCAGCCAT CTGTTGTTTG  
 GCGACTAGTC GGAGCTGACA CGGAAGATCA ACGGTCGGTA GACAACAAAC

2951 CCCCTCCCC GTGCCTTCCT TGACCCTGGA AGGTGCCACT CCCACTGTCC  
 GGGGAGGGGG CACGAAGGA ACTGGGACCT TCCACGGTGA GGGTGACAGG

3001 TTTCTTAATA AAATGAGGAA ATTGCATCGC ATTGTCTGAG TAGGTGTCAT  
 AAAGGATTAT TTTACTCCTT TAACGTAGCG TAACAGACTC ATCCACAGTA

3051 TCTATTCTGG GGGGTGGGGT GGGGCAGGAC AGCAAGGGGG AGGATTGGGA  
 AGATAAGACC CCCACCCCA CCCCGTCTG TCGTTCCCC TCCTAACCCCT

3101 AGACAATAGC AGGCATGCTG GGGAGCTCTT CCGCTTCCTC GCTCACTGAC  
 TCTGTTATCG TCCGTACGAC CCCTCGAGAA GGCGAAGGAG CGAGTGACTG

3151 TCGCTGCGCT CGGTCGTTCT GCTGCGGCGA GCGGTATCAG CTCACTCAA  
 AGCGACGCGA GCCAGCAAGC CGACGCCGCT CGCCATAGTC GAGTGAGTTT

Pci I

3201 GGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA GGAAAGAACA  
 CCGCCATTAT GCCAATAGGT GTCTTAGTCC CCTATTGCGT CCTTTCTTGT

Pci I

3251 TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG  
 AACTCGTTT TCCGGTCGTT TTCCGGTCCT TGGCATTITT CCGGCGCAAC

3301 CTGGCGTTTT TCCATAGGCT CCGCCCCCT GACGAGCATC AAAAAATCG  
 GACCGCAAAA AGGTATCCGA GCGGGGGGA CTGCTCGTAG TGTTTTTAGC

3351 ACGCTCAAGT CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG  
 TGCGAGTTCA GTCTCCACCG CTTTGGGCTG TCCTGATATT TCTATGGTCC

3401 CGTTTCCCC TGGAAGCTCC CTCGTGCGCT CTCCTGTTCC GACCCTGCCG  
 GCAAAGGGG ACCTTCGAGG GAGCACGCGA GAGGACAAGG CTGGGACGGC

FIG. 2F

HaeII

3451	CTTACCGGAT	ACCTGTCCGC	CTTTCTCCCT	TCGGGAAGCG	TGGCGCTTTC
	GAATGGCCTA	TGGACAGGCG	GAAAGAGGGA	AGCCCTTCGC	ACCGCGAAAG
<hr/>					
3501	TCAATGCTCA	CGCTGTAGGT	ATCTCAGTTC	GGTGTAGGTC	GTTGCTCCA
	AGTTACGAGT	GCGACATCCA	TAGAGTCAAG	CCACATCCAG	CAAGCGAGGT
<hr/>					
3551	AGCTGGGCTG	TGTGCACGAA	CCCCCGTTC	AGCCCGACCG	CTGCGCCTTA
	TCGACCCGAC	ACACGTGCTT	GGGGGGCAAG	TCGGGCTGGC	GACGCGGAAT
<hr/>					
3601	TCCGGTAACT	ATCGTCTTGA	GTCCAACCCG	GTAAGACACG	ACTTATCGCC
	AGGCCATTGA	TAGCAGAACT	CAGGTTGGGC	CATTCTGTGC	TGAATAGCGG
<hr/>					
3651	ACTGGCAGCA	GCCACTGGTA	ACAGGATTAG	CAGAGCGAGG	TATGTAGGCG
	TGACCGTCGT	CGGTGACCAT	TGTCCTAATC	GTCTCGCTCC	ATACATCCGC
<hr/>					
3701	GTGCTACAGA	GTTCTTGAAG	TGGTGGCCTA	ACTACGGCTA	CACTAGAAGG
	CACGATGTCT	CAAGAACTTC	ACCACCGGAT	TGATGCCGAT	GTGATCTTCC
<hr/>					
3751	ACAGTATTTG	GTATCTGCGC	TCTGCTGAAG	CCAGTTACCT	TCGGAAAAAG
	TGTCATAAAC	CATAGACGCG	AGACGACTTC	GGTCAATGGA	AGCCTTTTTTC
<hr/>					
3801	AGTTGGTAGC	TCTTGATCCG	GCAAACAAAC	CACCGCTGGT	AGCGGTGGTT
	TCAACCATCG	AGAACTAGGC	CGTTTGTTTG	GTGGCGACCA	TCGCCACCAA
<hr/>					
3851	TTTTTGTTTG	CAAGCAGCAG	ATTACGCGCA	GAAAAAAGG	ATCTCAAGAA
	AAAAACAAAC	GTTGCTGCTC	TAATGCGCGT	CTTTTTTTCC	TAGAGTTCTT
<hr/>					
3901	GATCCTTTGA	TCTTTTCTAC	GGGGTCTGAC	GCTCAGTGGA	ACGAAAACCT
	CTAGGAAACT	AGAAAAGATG	CCCCAGACTG	CGAGTCACCT	TGCTTTTGAG
<hr/>					
3951	ACGTTAAGGG	ATTTTGGTCA	TGAGATTATC	AAAAAGGATC	TTCACCTAGA
	TGCAATTCCC	TAAAACCACT	ACTCTAATAG	TTTTTCCTAG	AAGTGGATCT
<hr/>					
4001	TCCTTTTAAA	TTAAAAATGA	AGTTTAAAT	CAATCTAAAG	TATATATGAG
	AGGAAAATTT	AATTTTACT	TCAAAATTTA	GTTAGATTTC	ATATATACTC
<hr/>					
4051	TAAACTTGGT	CTGACAGTTA	CCAATGCTTA	ATCAGTGAGG	CACCTATCTC
	ATTTGAACCA	GACTGTCAAT	GGTTACGAAT	TAGTCACTCC	GTGGATAGAG

Eam1105I

AspEI

4101	AGCGATCTGT	CTATTTGCTT	CATCCATAGT	TGCCTGACTC	CCCGTCGTGT
	TCGCTAGACA	GATAAAGCAA	GTAGGTATCA	ACGGACTGAG	GGGCAGCACA
<hr/>					
4151	AGATAACTAC	GATACGGGAG	GGCTTACCAT	CTGGCCCCAG	TGCTGCAATG
	TCTATTGATG	CTATGCCCTC	CCGAATGGTA	GACCGGGGTC	ACGACGTTAC

FIG. 2G

Cfr101

~~~~~  
BsrFI  
~~~~~

4201 ATACCGCGAG ACCCAGGCTC ACCGGCTCCA GATTATCAG CAATAAACCA  
TATGGCGCTC TGGGTGCGAG TGGCCGAGGT CTAAATAGTC GTTATTGGT

BsaI  
~~~~~

4251 GCCAGCCGGA AGGGCCGAGC GCAGAAGTGG TCCTGCAACT TTATCCGCCT  
CGGTGCGCCT TCCCGGCTCG CGTCTTCACC AGGACGTTGA AATAGGCGGA

4301 CCATCCAGTC TATTAATTGT TGCCGGAAG CTAGAGTAAG TAGTTCGCCA  
GGTAGGTCAG ATAATTAACA ACGGCCCTTC GATCTCATTG ATCAAGCGGT

FspI  
~~~~~

AviII  
~~~~~

AosI  
~~~~~

4351 GTTAATAGTT TGCACAACGT TGTGCCATT GCTACAGGCA TCGTGGTGTG  
CAATTATCAA ACGCGTTGCA ACAACGGTAA CGATGTCCGT AGCACCACAG

4401 ACGCTCGTCG TTTGGTATGG CTTCAATCAG CTCCGGTTCC CAACGATCAA  
TGCGAGCAGC AAACCATACC GAAGTAAGTC GAGGCCAAGG GTTGCTAGTT

4451 GGCGAGTTAC ATGATCCCCC ATGTTGTGCA AAAAAGCGGT TAGCTCCTTC  
CCGCTCAATG TACTAGGGGG TACAACACGT TTTTTCGCCA ATCGAGGAAG

PvuI  
~~~~~

4501 GGTCTCCGA TCGTTGTCAG AAGTAAGTTG GCCGAGTGT TATCACTCAT  
CCAGGAGGCT AGCAACAGTC TTCATTCAAC CGGCGTCACA ATAGTGAGTA

4551 GGTTATGGCA GCACTGCATA ATTCTCTTAC TGTCATGCCA TCCGTAAGAT  
CCAATACCGT CGTGACGTAT TAAGAGAATG ACAGTACGGT AGGCATTCTA

4601 GCTTTTCTGT GACTGGTGAG TACTCAACCA AGTCATTCTG AGAATAGTGT  
CGAAAAGACA CTGACCACTC ATGAGTTGGT TCAGTAAGAC TCTTATCACA

BcgI  
~~~~~

4651 ATGCGGCGAC CGAGTTGCTC TTGCCCCGGC TCAATACGGG ATAATACCGC  
TACGCCGCTG GCTCAACGAG AACGGGCCGC AGTTATGCCC TATTATGGCG

XmnI  
~~~~~

Asp700  
~~~~~

4701 GCCACATAGC AGAACTTTAA AAGTGCTCAT CATTGGAAAA CGTTCTTCGG  
CGGTGTATCG TCTTGAATTT TTCACGAGTA GTAACCTTTT GCAAGAAGCC

4751 GGCGAAAACCT CTCAAGGATC TTACCGCTGT TGAGATCCAG TTCGATGTAA  
CCGCTTTTGA GAGTTCCTAG AATGGCGACA ACTCTAGGTC AAGCTACATT

4801 CCCACTCGTG CACCCAACTG ATCTTCAGCA TCTTTTACTT TCACCAGCGT  
GGGTGAGCAC GTGGGTTGAC TAGAAGTCGT AGAAAATGAA AGTGGTCGCA

FIG. 2H

4851 TTCTGGGTGA GCAAAAACAG GAAGGCAAAA TGCCGCAAAA AAGGGAATAA  
AAGACCCACT CGTTTTTGTC CTCCCGTTTT ACGGCGTTTT TTCCCTTATT

---

4901 GGGCGACACG GAAATGTTGA ATACTCATAC TCTTCCTTTT TCAATATTAT  
CCCGCTGTGC CTTTACAAC TATGAGTATG AGAAGGAAAA AGTTATAATA

---

4951 TGAAGCATT: ATCAGGGTTA TTGTCTCATG AGCGGATACA TATTGAATG  
ACTTCGTAA TAGTCCCAAT AACAGAGTAC TCGCCTATGT ATAAACTTAC

---

5001 TATTTAGAAA AATAAACAAA TAGGGGTTC GCGCACATTT CCCCAGAAAAG  
ATAAATCTTT TTATTTGTTT ATCCCAAGG CGCGTGTAAG GGGGCTTTTC

---

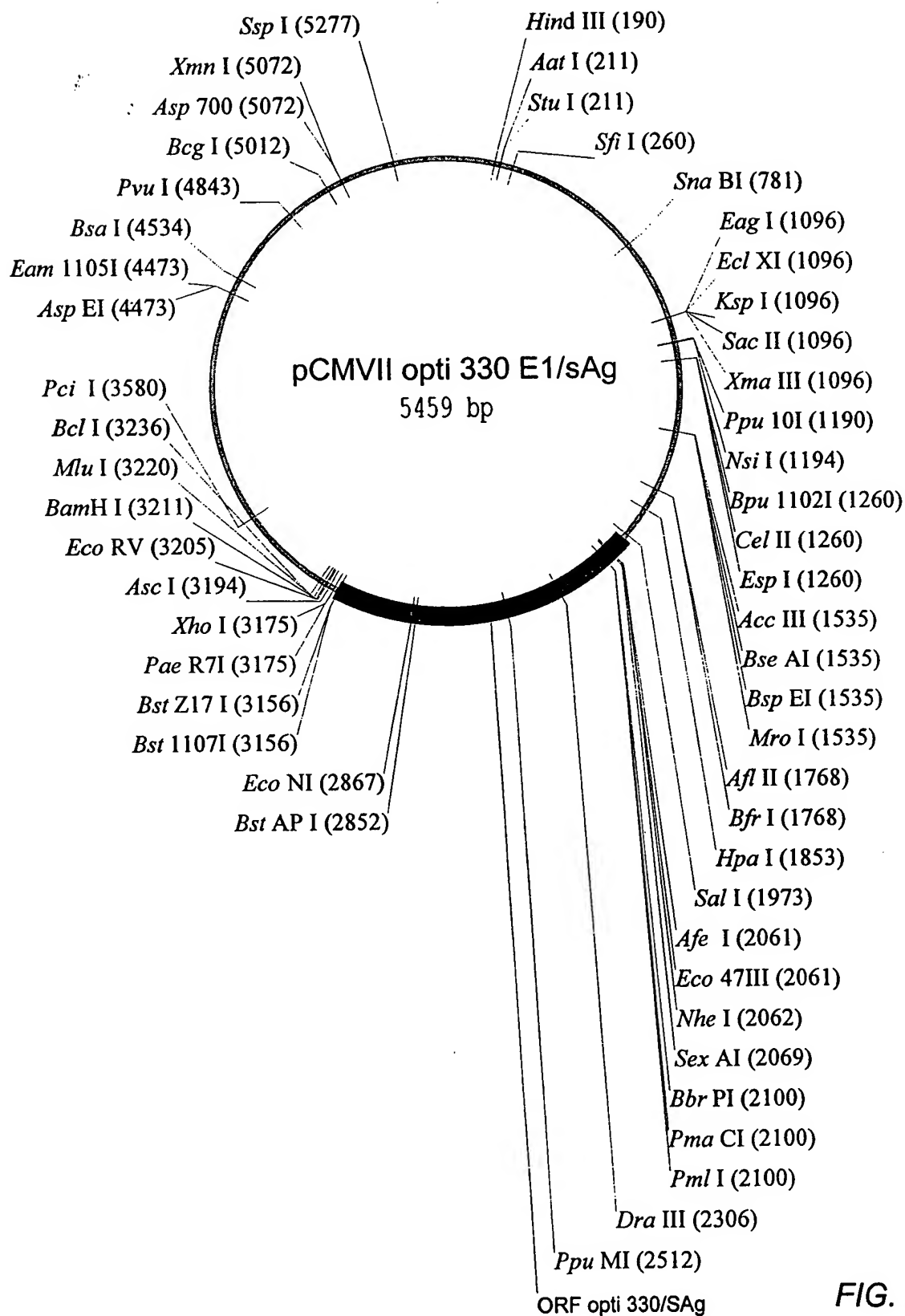
5051 TGCCACCTGA CGTCTAAGAA ACCATTATTA TCATGACATT AACCTATAAA  
ACGGTGGACT GCAGATTCTT TGGTAATAAT AGTACTGTAA TTGGATATT

---

5101 AATAGGCGTA TCACGAGGCC CTTTCGTC  
TTATCCGCAT AGTGCTCCGG GAAAGCAG

---

FIG. 2I



**FIG. 3A**



SEQ ID NO: 4

1 TCGCGCGTTT CGGTGATGAC TGAAGAAC TCTGACACAT GCAGCTCCCG  
AGCGCGCAAA GCCACTAC CCACTTTTGG AGACTGTGTA CGTCGAGGGC

---

51 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG  
CTCTGCCAGT GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTGCGGC

---

101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG  
AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCCGACC GAATTGATAC

---

HindIII  
~~~~~

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGTCA  
GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT

---

StuI  
~~~~~  
AatI  
~~~~~

201 AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG  
TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

---

SfiI  
~~~~~

251 AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA  
TCCGGCTCCG CCGGAGCCGG AGACGTATT ATTTTTTTTA ATCAGTCGGT

---

301 TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG  
ACCCCGCCTC TTACCCGCCT TGACCCGCC CTCCCTTAAT AACCGATAAC

---

351 GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

---

401 CATGTCCAAT ATGACCGCCA TGTGACATT GATTATTGAC TAGTTATTAA  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

---

451 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTCCG  
ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

---

501 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC  
GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

---

551 CCCGCCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA  
GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTAT

---

601 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AACTGCCCCA  
CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

---

651 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG  
GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

---

701 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA  
AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

---

SnaBI  
~~~~~

751 CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

FIG. 3B

801 CATGGTGATG CGGTTT . AGTACACCAA TGGGCGTGGA TAGCGGTTTG  
GTACCACTAC GCCAAAACCG TCATGTGGT ACCCGCACCT ATCGCCAAAC

851 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG  
TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC

901 TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC  
AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG

951 CCCGTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA  
GGGCAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT

1001 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC  
CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG

XmaIII

SacII

KspI

EclXI

EagI

1051 TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG  
ACAAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC

1101 GGAACGGTGC ATTGGAACGC GGATTCCCG TGCCAAGAGT GACGTAAGTA  
CCTTGCCACG TAACCTTGCG CTAAGGGGC ACGGTCTCA CTGCATTCAT

Ppu10I

NsiI

1151 CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCG GCTCCTTATG CTATAGGTGA  
GACAAAACCG GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT

EspI

CelII

Bpu1102I

1251 TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC  
ACCATATCGA ATCGGATATC CACACCCAAT AACTGGTAAT AACTGGTGAG

1301 CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG  
GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC

1351 CCACAACAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT  
GGTGTTGATA GAGATAACCG ATATACGGT ATGAGACAGG AAGTCTCTGA

1401 GACACGGACT CTGTATTTT ACAGGATGGG GTCCATTTAT TATTACAAA  
CTGTGCCTGA GACATAAAAA TGTCTACCC CAGGTAAATA ATAAATGTTT

FIG. 3C

1451 TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA  
AAGTGATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTGT

MroI

BspEI

BseAI

AccIII

1501 TAGCGTGGGA TCTCCGACAT CTCGGGTACG TGTTCCGGAC ATGGGCTCTT  
ATCGCACCT AGAGGCTGTA GAGCCCATGC ACAAGGCCTG TACCCGAGAA

1551 CTCCGGTAGC GGCAGGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
GAGGCCATCG CCGCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

1651 CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT  
GAATCCGTGT CGTGTACGG GTGGTGGTGG TCACACGGCG TGTTCCGGCA

1701 GCGGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT  
CGGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA

BfrI

AflII

1751 GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT  
CCTGCGTCTA CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA

HpaI

1801 GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TGCGGTGCTG  
CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC

HpaI

1851 TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG  
AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

1901 CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG  
GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

+3

SEQ ID No: 5 → M D A

SalI

1951 GTCTTTTCTG CAGTCACCGT CGTCGACGAA TTCAAGCAAT CATGGATGCA  
CAGAAAAGAC GTCAGTGGCA GCAGCTGCTT AAGTTCGTTA GTACCTACGT

+3 M K R G L C C V L L L C G A V F V  
2001 ATGAAGAGAG GGCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCGT  
TACTTCTCTC CCGAGACGAC ACACGACGAC GACACACCTC GTCAGAAGCA

FIG. 3D

```

+3  S P S  A S Y Q V R N S T G L Y H
      NheI
      -----
      Eco47III
      -----
      Afe I      SexAI
      -----
      PmlI
      -----
      PmaCI
      -----
      BbrPI
      -----
2051  TTCGCCCAGC GCTAGCTACC AGGTGCGCAA CAGCACCAGC CTGTACCACG
      AAGCGGCTCG CGATCGATGG TCCACGCGTT GTCGTGGCCG GACATGGTGC

+3  V T N D C P N S S I V Y E A A D A
      PmlI
      --
      PmaCI
      --
      BbrPI
      --
2101  TGACCAACGA CTGCCCCAAC AGCAGCATCG TGTACGAGGC CGCCGACGCC
      ACTGGTTGCT GACGGGGTTG TCGTCGTAGC ACATGCTCCG GCGGCTGCGG

+3  I L H T P G C V P C V R E G N A S
2151  ATCCTGCACA CCCCCGGCTG CGTGCCCTGC GTGCGCGAGG GCAACGCCAG
      TAGGACGTGT GGGGGCCGAC GCACGGGACG CACGCGCTCC CGTGCGGTGC

+3  R C W V A M T P T V A T R D G K
2201  CCGCTGCTGG GTGGCCATGA CCCCCACCGT GGCCACCCGC GACGGCAAGC
      GCGCAGCACC CACCGGTACT GGGGGTGGCA CCGGTGGGCG CTGCCGTTGC

+3  L P A T Q L R R H I D L L V G S A
      DraIII
      ~
2251  TGCCCGCCAC CCAGCTGCGC CGCCACATCG ACCTGCTGGT GGGCAGCGCC
      ACGGGCGGTG GGTGACGCGG GCGGTGTAGC TGGACGACCA CCGTCGCGG

+3  T L C S A L Y V G D L C G S V F L
      DraIII
      -----
2301  ACCCTGTGCA GCGCCCTGTA CGTGGGCGAC CTGTGCGGCA GCGTGTTCCT
      TGGGACACGT CGCGGGACAT GCACCCGCTG GACACGCCGT CGCACAAGGA

+3  V G Q L F T F S P R R H W T T Q
2351  GGTGGGCCAG CTGTTACCTC TCAGCCCCCG CCGCCACTGG ACCACCCAGG
      CCACCCGGTC GACAAGTGGG AGTCGGGGGC GCGGTGACC TGGTGGGTCC

+3  G C N C S I Y P G H I T G H R M A
2401  GCTGCAACTG CAGCATCTAC CCGGCCACA TCACCGCCA CCGCATGGCC
      CGACGTTGAC GTCGTAGATG GGGCCGGTGT AGTGGCCGGT GGCGTACCGG

+3  W D M M M N W S P T T M E N I T S
2451  TGGGACATGA TGATGAAGT GAGCCCCACC ACCATGGAGA ACATCACATC
      ACCCTGTACT ACTACTTGAC CTCGGGGTGG TGGTACCTCT TGTAGTGTAG

+3  G F L G P L L V L Q A G F F L L
      PpuMI
      -----
2501  AGGATTCTTA GGACCCCTGC TCGTGTTACA GGCGGGGTTT TTCTTGTTGA
      TCCTAAGGAT CCTGGGGACG AGCACAATGT CCGCCCCAAA AAGAACAAC

```

FIG. 3E

+3 T R I L . T I P Q S L D S W W T S L  
 2551 CAAGAATCCT CACAATACCG CAGAGTCTAG ACTCGTGGTG GACTTCTCTC  
 GTTCTTAGGA GTGTTATGGC GTCTCAGATC TGAGCACCAC CTGAAGAGAG

---

+3 N F L G G S P V C L G Q N S Q S P  
 2601 AATTTTCTAG GGGGATCTCC CGTGTGTCTT GGCCAAAATT CGCAGTCCCC  
 TTAAAAGATC CCCCTAGAGG GCACACAGAA CCGGTTTAA GCGTCAGGGG

---

+3 T S N H S P T S C P P I C P G Y  
 2651 AACCTCCAAT CACTCACCAA CCTCCTGTCC TCCAATTTGT CCTGGTTATC  
 TTGGAGGTTA GTGAGTGGTT GGAGGACAGG AGGTAAACA GGACCAATAG

---

+3 R W M C L R R F I I F L F I L L L  
 2701 GCTGGATGTG TCTGCGGCGT TTTATCATAT TCCTCTTCAT CCTGCTGCTA  
 CGACCTACAC AGACGCCGCA AAATAGTATA AGGAGAAGTA GGACGACGAT

---

+3 C L I F L L V L L D Y Q G M L P V  
 2751 TGCCTCATCT TCTTATTGGT TCTTCTGGAT TATCAAGGTA TGTTGCCCGT  
 ACGGAGTAGA AGAATAACCA AGAAGACCTA ATAGTTCCAT ACAACGGGCA

---

+3 C P L I P G S T T T S T G P C K  
 BstAP I  
 2801 TTGTCCTCTA ATTCCAGGAT CAACAACAAC CAGTACGGGA CCATGCAAAA  
 AACAGGAGAT TAAGGTCCTA GTTGTGTGTG GTCATGCCCT GGTACGTTTT

---

+3 T C T T P A Q G N S M F P S C C C  
 BstAP I EcoNI  
 2851 CCTGCACGAC TCCTGTCTCA GGCAACTCTA TGTTTCCCTC ATGTTGCTGT  
 GGACGTGCTG AGGACGAGTT CCGTTGAGAT ACAAAGGGAG TACAACGACA

---

+3 T K P T D G N C T C I P I P S S W  
 2901 AAAAAACCTA CGGATGGAAA TTGCACCTGT ATTCCCATCC CATCGTCTG  
 TGTTTTGGAT GCCTACCTTT AACGTGGACA TAAGGGTAGG GTAGCAGGAC

---

+3 A F A K Y L W E W A S V R F S W  
 2951 GGCTTTTCGCA AAATACCTAT GGGAGTGGGC CTCAGTCCGT TTCTCTTGGC  
 CCGAAAGCGT TTTATGGATA CCCTCACCCG GAGTCAGGCA AAGAGAACCG

---

+3 L S L L V P F V Q W F V G L S P T  
 3001 TCAGTTTACT AGTGCCATTT GTTCAGTGGT TCGTAGGGCT TTCCCCACT  
 AGTCAAATGA TCACGGTAAA CAAGTCACCA AGCATCCCGA AAGGGGGTGA

---

+3 V W L S A I W M M W Y W G P S L Y  
 3051 GTTTGGCTTT CAGCTATATG GATGATGTGG TATTGGGGGC CAAGTCTGTA  
 CAAACCGAAA GTCGATATAC CTACTACACC ATAACCCCGG GTTCAGACAT

---

+3 S I V S P F I P L L P I F F C L  
 3101 CAGCATCGTG AGTCCCTTTA TACCGCTGTT ACCAATTTTC TTTGTCTCT  
 GTCGTAGCAC TCAGGGAAAT ATGGCGACAA TGGTTAAAG AAAACAGAGA

---

+3 W V Y I \*  
 BstZ17 I XhoI  
 Bst1107I PaeR7I AscI  
 3151 GGGTATACAT TTAAGAATTC AGACTCGAGC AAGTCTAGAA AGGCGCGCCA  
 CCCATATGTA AATTCTTAAG TCTGAGCTCG TTCAGATCTT TCCGCGCGGT

FIG. 3F

|      | EcoRV                    | BamHI                     | MluI                     | BclI                     |                          |
|------|--------------------------|---------------------------|--------------------------|--------------------------|--------------------------|
| 3201 | AGATATCAAG<br>TCTATAGTTC | GATCCACTAC<br>CTAGGTGATG  | CGGTTAGAGC<br>CGCAATCTCG | TCGCTGATCA<br>AGCGACTAGT | GCCTCGACTG<br>CGGAGCTGAC |
| 3251 | TGCCTTCTAG<br>ACGGAAGATC | TTGCCAGCCA<br>AACGGTCGGT  | TCTGTTGTTT<br>AGACAACAAA | CCCCCTCCCC<br>CGGGGAGGGG | CGTGCCTTCC<br>GCACGGAAG  |
| 3301 | TTGACCCTGG<br>AACTGGGACC | AAGGTGCCAC<br>TTCCACGGTG  | TCCCACTGTC<br>AGGGTGACAG | CTTTCCTAAT<br>GAAAGGATTA | AAAATGAGGA<br>TTTTACTCCT |
| 3351 | AATTGCATCG<br>TTAACGTAGC | CATTGTCTGA<br>GTAACAGACT  | GTAGGTGTCA<br>CATCCACAGT | TTCTATTCTG<br>AAGATAAGAC | GGGGGTGGGG<br>CCCCCACCCC |
| 3401 | TGGGGCAGGA<br>ACCCCGTCCT | CAGCAAGGGG<br>GTCGTTCCCC  | GAGGATTGGG<br>CTCCTAACCC | AAGACAATAG<br>TTCTGTTATC | CAGGCATGCT<br>GTCCGTACGA |
| 3451 | GGGGAGCTCT<br>CCCCTCGAGA | TCCGCTTCCT<br>AGGCGAAGGA  | CGCTCACTGA<br>GCGAGTGACT | CTCGCTGCGC<br>GAGCGACGCG | TCGGTCGTTT<br>AGCCAGCAAG |
| 3501 | GGCTGCGGCG<br>CCGACGCCGC | AGCGGTATCA<br>TCGCCATAGT  | GCTCACTCAA<br>CGAGTGAGTT | AGGCGGTAAT<br>TCCGCCATTA | ACGGTTATCC<br>TGCCAATAGG |
|      |                          |                           | Pci I                    |                          |                          |
| 3551 | ACAGAATCAG<br>TGTCTTAGTC | GGGATAACGC<br>CCCTATTGCG  | AGGAAAGAAC<br>TCCTTTCTTG | ATGTGAGCAA<br>TACACTCGTT | AAGGCCAGCA<br>TTCCGGTCGT |
| 3601 | AAAGGCCAGG<br>TTTCCGGTCC | AACCGTAAAA<br>TTGGCATTTC  | AGGCCGCGTT<br>TCCGGCGCAA | GCTGGCGTTT<br>CGACCGCAA  | TTCCATAGGC<br>AAGGTATCCG |
| 3651 | TCCGCCCCCC<br>AGGCGGGGGG | TGACGAGCAT<br>ACTGCTCGTA  | CACAAAAATC<br>GTGTTTTTAG | GACGCTCAAG<br>CTGCGAGTTC | TCAGAGGTGG<br>AGTCTCCACC |
| 3701 | CGAAACCCGA<br>GCTTTGGGCT | CAGGACTATA<br>GTCCTGATAT  | AAGATACCAG<br>TTCTATGGTC | GCGTTTCCCC<br>CGCAAAGGGG | CTGGAAGCTC<br>GACCTTCGAG |
| 3751 | CCTCGTGCGC<br>GGAGCACGCG | TCTCCTGTTC<br>AGAGGACAAG  | CGACCCTGCC<br>GCTGGGACGG | GCTTACCGGA<br>CGAATGGCCT | TACCTGTCCG<br>ATGGACAGGC |
| 3801 | CCTTTCTCCC<br>GGAAAGAGGG | TTCCGGGAAGC<br>AAGCCCTTCG | GTGGCGCTTT<br>CACCGCGAAA | CTCAATGCTC<br>GAGTTACGAG | ACGCTGTAGG<br>TGCACATCC  |
| 3851 | TATCTCAGTT<br>ATAGAGTCAA | CGGTGTAGGT<br>GCCACATCCA  | CGTTCGCTCC<br>GCAAGCGAGG | AAGCTGGGCT<br>TTCGACCCGA | GTGTGCACGA<br>CACACGTGCT |
| 3901 | ACCCCCCGTT<br>TGGGGGGCAA | CAGCCCGACC<br>GTCGGGCTGG  | GCTGCGCCTT<br>CGACGCGGAA | ATCCGGTAAC<br>TAGGCCATTG | TATCGTCTTG<br>ATAGCAGAAC |
| 3951 | AGTCCAACCC<br>TCAGGTTGGG | GGTAAGACAC<br>CCATTCTGTG  | GACTTATCGC<br>CTGAATAGCG | CACTGGCAGC<br>GTGACCGTCG | AGCCACTGGT<br>TCGGTGACCA |
| 4001 | AACAGGATTA<br>TTGTCCTAAT | GCAGAGCGAG<br>CGTCTCGCTC  | GTATGTAGGC<br>CATAATCCG  | GGTGCTACAG<br>CCACGATGTC | AGTTCTTGAA<br>TCAAGAATT  |
| 4051 | GTGGTGGCCT<br>CACCACCGGA | AACTACGGCT<br>TTGATGCCGA  | ACACTAGAAG<br>TGTGATCTTC | GACAGTATTT<br>CTGTCATAAA | GGTATCTGCG<br>CCATAGACGC |

FIG. 3G

4101 CTCTGCTGAA GCC ACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC  
GAGACGACTT CGGTCAATGG AAGCCTTTTT CTCAACCATC GAGAACTAGG

4151 GGCAAACAAA CCACCGCTGG TAGCGGTGGT TTTTTGTTT GCAAGCAGCA  
CCGTTTGTTT GGTGGCGACC ATCGCCACCA AAAAAACAAA CGTTCGTCGT

4201 GATTACGCGC AGAAAAAAG GATCTCAAGA AGATCCTTTG ATCTTTTCTA  
CTAATGCGCG TCTTTTTTTC CTAGAGTTCT TCTAGGAAAC TAGAAAAGAT

4251 CGGGGTCTGA CGCTCAGTGG AACGAAACT CACGTTAAGG GATTTTGGTC  
GCCCCAGACT GCGAGTCACC TTGCTTTTGA GTGCAATTCC CTAAAACCAG

4301 ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA ATTAAAAATG  
TACTCTAATA GTTTTTCCTA GAAGTGGATC TAGGAAAATT TAATTTTAC

4351 AAGTTTAAA TCAATCTAAA GTATATATGA GTAACTTGG TCTGACAGTT  
TTCAAAATTT AGTTAGATT CATATATACT CATTGAACC AGACTGTCAA

4401 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCGT  
TGTTACGAA TTAGTCACTC CGTGGATAGA GTCGCTAGAC AGATAAGCA

Eam1105I

AspEI

4451 TCATCCATAG TTGCCTGACT CCCCCTCGTG TAGATAACTA CGATACGGGA  
AGTAGGTATC AACGGACTGA GGGGCAGCAC ATCTATTGAT GCTATGCCCT

4501 GGGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA GACCCACGCT  
CCCGAATGGT AGACCGGGGT CACGACGTTA CTATGGCGCT CTGGGTGCGA

BsaI

4551 CACCGGCTCC AGATTTATCA GCAATAAACC AGCCAGCCGG AAGGGCCGAG  
GTGGCCGAGG TCTAAATAGT CGTTATTTGG TCGGTCGGCC TTCCCGGCTC

4601 CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT CTATTAATTG  
GCGTCTTCAC CAGGACGTTG AAATAGGCGG AGGTAGGTCA GATAATTAAC

4651 TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG  
AACGGCCCTT CGATCTCATT CATCAAGCGG TCAATTATCA AACGCGTTGC

4701 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG  
AACAAACGGTA ACGATGTCCG TAGCACCACA GTGCGAGCAG CAAACCATAC

4751 GCTTCATTCA GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC  
CGAAGTAAGT CGAGGCCAAG GGTGCTAGT TCCGCTCAAT GTACTAGGGG

PvuI

4801 CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA  
GTACAACAG TTTTTTCGCC AATCGAGGAA GCCAGGAGGC TAGCAACAGT

4851 GAAGTAAGTT GGCCGCGAGTG TTATCACTCA TGGTTATGGC AGCACTGCAT  
CTTCATTCAA CCGGCGTCAC AATAGTGAGT ACCAATACCG TCGTGACGTA

4901 AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG TGA CTGGTGA  
TTAAGAGAAT GACAGTACGG TAGGCATTCT ACGAAAGAC ACTGACCACT

FIG. 3H

BcgI  
-----

4951 GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT  
CATGAGTTGG TTCAGTAAGA CTCTTATCAC ATACGCCGCT GGCTCAACGA

---

5001 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA  
GAACGGGGCG CAGTTATGCC CTATTATGGC GCGGTGTATC GTCTTGAAAT

---

XmnI  
-----  
Asp700  
-----

5051 AAAGTGCTCA TCATTGGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT  
TTTCACGAGT AGTAACCTTT TGCAAGAAGC CCCGCTTTG AGAGTTCCTA

---

5101 CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT GCACCCAACT  
GAATGGCGAC AACTCTAGGT CAAGCTACAT TGGGTGAGCA CGTGGGTTGA

---

5151 GATCTTCAGC ATCTTTTACT TTCACCAGCG TTTCTGGGTG AGCAAAAACA  
CTAGAAGTCG TAGAAAATGA AAGTGGTCGC AAAGACCCAC TCGTTTTGT

---

5201 GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC GGAAATGTTG  
CCTTCCGTTT TACGGCGTTT TTTCCCTTAT TCCCGCTGTG CCTTTACAAC

---

SspI  
-----

5251 AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT  
TTATGAGTAT GAGAAGGAAA AAGTTATAAT AACTTCGTAA ATAGTCCCAA

---

5301 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA  
TAACAGAGTA CTCGCCTATG TATAAACTTA CATAAATCTT TTTATTTGTT

---

5351 ATAGGGGTTT CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA  
TATCCCCAAG GCGCGTGTA AGGGGCTTTT CACGGTGGAC TGCAGATTCT

---

5401 AACCATTATT ATCATGACAT TAACCTATAA AAATAGGCGT ATCAGAGGC  
TTGGTAATAA TAGTACTGTA ATTGGATATT TTTATCCGCA TAGTGCTCCG

---

5451 CCTTTCGTC  
GGAAAGCAG

---

FIG. 3I



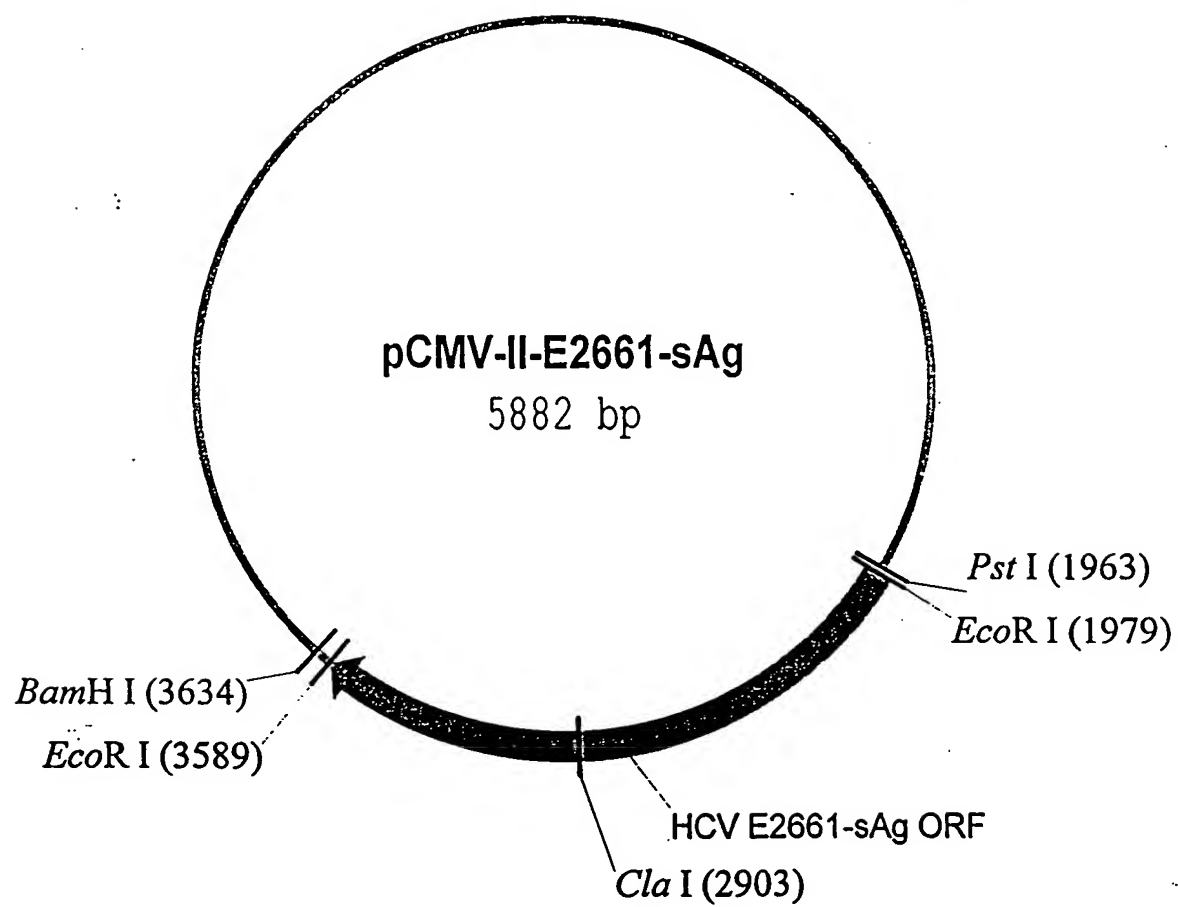


FIG. 4A

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT  
AGCGCGCAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTGCGCTA

81 GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA  
CGGCCCTCGT CTGTTCCGGC AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT

161 GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA AAAGCCTAGG CCTCCAAAA AGCCTCCTCA CTACTTCTGG  
CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT TTTCCGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC

241 AATAGCTCAG AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGGCGGA  
TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT ACCCCGCCCTC TTACCCGCCT

321 ACTGGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
TGACCCGCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401 CATGTCCAAT ATGACCGCCA TGTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA

481 AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCCCG TGGCTGACCG CCCAACGACC CCCGCCATT  
TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA

561 GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT  
CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

641 AAAGTGGCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG TCAATGACGG TAAATGGCCC  
TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC AGTTACTGCC ATTTACCGGG

721 GCCTGGCATT ATGCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
CGGACCGTAA TACGGGTCAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

801 CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA  
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGTGCGGGT

881 TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC CCCGTTGACG  
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

961 CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG  
GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGCGCG GGAACGGTGC ATTGGAACGC  
GGTAGGTGCG ACAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGCG AGGCGCCGCG CCTTGCCACG TAACCTTGCG

1121 GGATTCCCCG TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
CCTAAGGGGC ACGGTTCTCA CTGCATTCTT GCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCG GTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA  
GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

1281 TTGACCATTA TTGACCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACAT  
AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

1361 CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT  
GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCTACCC CAGGTAAATA

1441 TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCCGAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT  
ATAAATGTTT AAGTGATATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAA AATAATTTGT ATCGCACCTC AGAGGCTGTA

FIG. 4B

1521 CTCGGGTACG TGTTCGGAC ATGGGCTCTT CTCCGGTAGC GCGGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
GAGCCCATGC ACAAGGCCGTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGGCTCATG GTCGCTCGGC AGTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACGG GTGGTGGTGG

1681 AGTGTGCCGC ACAAGGCCGT GCGGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT  
TCACACGGCG TGTTCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGA CCTGCGTCTA

1761 GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCCT  
CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA CTCACAAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841 TCGGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT  
ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

SEQ ID NO: 7 → M D A

+3 PstI EcoRI  
~~~~~

1921 GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACGAA TTCAAGCAAT CATGGATGCA  
CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGCTT AAGTTCGTTA GTACCTACGT

+3 M K R G L C C V L L L C G A V F V S P S A S E T H V T  
2001 ATGAAGAGAG GGCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCGT TTCGCCAGC GCTAGCGAAA CCCACGTCAC  
TACTTCTCTC CCGAGACGAC ACACGACGAC GACACACCTC GTCAGAAGCA AAGCGGGTCG CGATCGCTTT GGGTGCAGTG

+3 G G S A G H T V S G F V S L L A P G A K Q N V Q L I  
2081 CGGGGGAAGT GCCGGCCACA CTGTGTCTGG ATTTGTTAGC CTCCTCGCAC CAGGCGCCAA GCAGAACGTC CAGCTGATCA  
GCCCCCTTCA CGGCCGTGT GACACAGACC TAAACAATCG GAGGAGCGTG GTCCGCGGTT CGTCTGTCAG GTCGACTAGT

+3 N T N G S W H L N S T A L N C N D S L N T G W L A G L  
2161 ACACCAACGG CAGTTGGCAC CTCAATAGCA CGGCCCTGAA CTGCAATGAT AGCCTCAACA CCGCTGGTT GGCAGGGCTT  
TGTGGTTGCC GTCACCGTG GAGTTATCGT GCCGGGACTT GACGTTACTA TCGGAGTTGT GGCCGACCA CCGTCCCGAA

+3 F Y H H K F N S S G C P E R L A S C R P L T D F D Q G  
2241 TTCTATCACC ACAAGTTCAA CTCTTCAGGC TGTCCTGAGA GGCTAGCCAG CTGCCGACCC CTTACCGATT TTGACCAGG  
AAGATAGTGG TGTTCAGTT GAGAAGTCCG ACAGGACTCT CCGATCGGTC GACGGCTGGG GAATGGCTAA AACTGGTCCC

+3 W G P I S Y A N G S G P D Q R P Y C W H Y P P K P C  
2321 CTGGGGCCCT ATCAGTTATG CCAACGGAAG CGGCCCGGAC CAGCGCCCT ACTGCTGGCA CTACCCCCCA AAACCTTGCG  
GACCCCGGGA TAGTCAATAC GGTGCGCTTC GCCGGGGCTG GTCGCGGGGA TGACGACCGT GATGGGGGGT TTTGGAACGC

+3 G I V P A K S V C G P V Y C F T P S P V V V G T T D R  
2401 GTATTGTGCC CGCGAAGAGT GTGTGTGGTC CCGTATATTG CTTCACTCCC AGCCCCGTGG TGGTGGGAAC GACCGACAGG  
CATAACACGG GCGTTTCTCA CACACACCAG GCCATATAAC GAAGTGAGGG TCGGGGCACC ACCACCCTTG CTGGCTGTCC

+3 S G A P T Y S W G E N D T D V F V L N N T R P P L G N  
2481 TCGGGCGCGC CCACCTACAG CTGGGGTGAA AATGATACGG ACGTCTTCGT CCTTAACAAT ACCAGGCCAC CGCTGGGCAA  
AGCCCGCGCG GGTGGATGTC GACCCCACTT TTAATATGCC TGCAGAAGCA GGAATTGTTA TGGTCCGGTG GCGACCCGTT

+3 W F G C T W M N S T G F T K V C G A P P C V I G G A  
2561 TTGGTTCGGT TGTACCTGGA TGAATCAAC TGGATTACCC AAAGTGTGCG GAGCGCCTCC TTGTGTCATC GGAGGGGCGG  
AACCAAGCCA ACATGGACCT ACTGAGTTG ACCTAAGTGG TTTACACGC CTCGCGGAGG AACACAGTAG CCTCCCCGCC

+3 G N N T L H C P T D C F R K H P D A T Y S R C G S G P  
2641 GCAACAACAC CCTGCACTGC CCCACTGATT GCTTCCGCAA GCATCCGGAC GCCACATACT CTCGGTGGCG CTCCGGTCCC  
CGTTGTTGTG GGACGTGACG GGGTGACTAA CGAAGGCGTT CGTAGGCCTG CCGTGTATGA GAGCCACGCC GAGGCCAGGG

FIG. 4C

+3 W I T P R C L V D Y P Y R L W H Y P C T I N Y T I F K  
 2721 TGGATCACAC CCAGGTGCCT GGTGACTAC CCGTATAGGC TTTGGCATTG TCCTTGTTACC ATCAACTACA CCATATTTAA  
 ACCTAGTGTG GGTCCACGGA CCAGCTGATG GGCATATCCG AAACCGTAAT AGGAACATGG TAGTTGATGT GGTATAAATT

+3 I R M Y V G G V E H R L E A A C N W T R G E R C D L  
 2801 AATCAGGATG TACGTGGGAG GGGTCGAACA CAGGCTGGAA GCTGCCTGCA ACTGGACGCG GGGCGAACGT TGCAGTCTGG  
 TTAGTCCTAC ATSCACCCTC CCCAGCTTGT GTCCGACCTT CGACGGACGT TGACCTGCGC CCCGCTTGCA ACGCTAGACC

+3 E D R D R S E I D M E N I T S G F L G P L L V L Q A G  
 ClaI  
 ~~~~~  
 2881 AAGATAGGGA CAGGTCCGAG ATCGATATGG AGAACATCAC ATCAGGATTC CTAGGACCCC TGCTCGTGTG ACAGGCGGGG  
 TTCTATCCCT GTCCAGGCTC TAGCTATACC TCTTGATAGT TAGTCTTAAG GATCCTGGGG ACGAGCACAA TGTCCGCCCC

+3 F F L L T R I L T I P Q S L D S W W T S L N F L G G S  
 2961 TTTTCTTGT TGACAAGAA CTTCACAATA CCGCAGAGTC TAGACTCGTG GTGGACTTCT CTCAATTTTC TAGGGGGATC  
 AAAAAGAACA ACTGTTCTTA GGAGTGTAT GCGTCTCAG ATCTGAGCAC CACCTGAAGA GAGTTAAAG ATCCCCTAG

+3 P V C L G Q N S Q S P T S N H S P T S C P P I C P G  
 3041 TCCCGTGTGT CTGGGCCAAA ATTCGCAGTC CCAAACCTCC AATCACTCAC CAACCTCCTG TCCTCCAATT TGCTCTGGTT  
 AGGGCACACA GAACCGGTTT TAAGCGTCAG GGGTTGGAGG TTAGTGAGTG GTTGGAGGAC AGGAGGTTAA ACAGGACCAA

+3 Y R W M C L R R F I I F L F I L L L C L I F L L V L L  
 3121 ATCGCTGGAT GTGTCTGCGG CGTTTTATCA TATTCTCTT CATCTGCTG CTATGCCTCA TCTTCTTATT GGTCTTCTG  
 TAGCGACCTA CACAGACGCC GCAAATAGT ATAAGGAGAA GTAGGACGAC GATACGGAGT AGAAGAATAA CCAAGAAGAC

+3 D Y Q G M L P V C P L I P G S T T T S T G P C K T C T  
 3201 GATTATCAAG GTATGTTGCC CGTTTGTCTT CTAATTCCAG GATCAACAAC AACCAGTACG GGACCATGCA AAACCTGCAC  
 CTAATAGTTC CATAACCGG GCAAACAGGA GATTAAGGTC CTAGTTGTTG TTGGTCATGC CCTGGTACGT TTTGGACGTG

+3 T P A Q G N S M F P S C C C T K P T D G N C T C I P  
 3281 GACTCCTGCT CAAGGCAACT CTATGTTTCC CTCATGTTGC TGTACAAAAC CTACGGATGG AAATTGCACC TGTATTCCCA  
 CTGAGGACGA GTTCCGTTGA GATACAAAGG GAGTACAACG ACATGTTTGG GATGCCTACC TTTACGTGG ACATAAGGGT

+3 I P S S W A F A K Y L W E W A S V R F S W L S L L V P  
 3361 TCCCATCGTC CTGGGCTTTC GCAAATACC TATGGGAGTG GGCCTCAGTC CGTTTCTCTT GGCTCAGTTT ACTAGTGCCA  
 AGGGTAGCAG GACCCGAAAG CGTTTTATGG ATACCCTCAC CCGGAGTCAG GCAAAGAGAA CCGAGTCAAA TGATCACGGT

+3 F V Q W F V G L S P T V W L S A I W M M W Y W G P S L  
 3441 TTTGTTTCA GTTTCGTAGG GCTTTCCCC ACTGTTTGGC TTTAGCTAT ATGGATGATG TGGTATTGGG GGCCAAGTCT  
 AAACAAGTCA CCAAGCATCC CGAAAGGGGG TGACAAACCG AAAGTCGATA TACCTACTAC ACCATAACCC CCGGTTTCA

+3 Y S I V S P F I P L L P I F F C L W V Y I \*  
 EcoRI  
 ~~~~~  
 3521 GTACAGCATC GTGAGTCCCT TTATACCGCT GTTACCAATT TTCTTTTGTG TCTGGGTATA CATTTAAGAA TTCAGACTCG  
 CATGTCGTAG CACTCAGGGA AATATGGCGA CAATGGTTAA AAGAAAACAG AGACCCATAT GTAAATTCTT AAGTCTGAGC

BamHI  
 ~~~~~  
 3601 AGCAAGTCTA GAAAGGCGCG CCAAGATATC AAGGATCCAC TACGCGTTAG AGCTCGCTGA TCAGCCTCGA CTGTGCCTTC  
 TCGTTCAGAT CTTTCCGCGC GGTCTATAG TTCTAGGTG ATGCGCAATC TCGAGCGACT AGTCGGAGCT GACACGGAAG

3681 TAGTTGCCAG CCATCTGTTG TTTGCCCCCT CCCCCTGCCT TCCTTGACCC TGGAAAGGTGC CACTCCCACT GTCCTTTTCT  
 ATCAACGGTC GGTAGACAAC AAACGGGGAG GGGGACGGA AGGAACTGGG ACCTTCCACG GTGAGGGTGA CAGGAAAGGA

3761 AATAAAATGA GGAAATTGCA TCGCATTGTC TGAGTAGGTG TCATTCTATT CTGGGGGGTG GGGTGGGGCA GGACAGCAAG  
 TTATTTTACT CCTTTAACGT AGCGTAACAG ACTCATCCAC AGTAAGATAA GACCCCCAC CCCACCCCGT CCTGTCGTTT

FIG. 4D

3841 GGGGAGGATT GGGAAGACAA TAGCAGGCAT GCTGGGGAGC TCTTCCGCTT CCTCGCTCAC TGA CTGCTG CGCTCGGTG  
 CCCCTCCTAA CCCTTCTGTT ATCGTCCGTA CGACCCCTCG AGAAGGCGAA GGAGCGAGTG ACTGAGCGAC GCGAGCCAGC

3921 TTCGGCTGCG GCGAGCGGTA TCAGCTCACT CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG  
 AAGCCGACGC CGCTCGCCAT AGTCGAGTGA GTTTCCGCCA TTATGCCAAT AGGTGTCTTA GTCCCTATT GCGTCTTTTC

4001 AACATGTGAG CAAAAGGCCA GAAAAGGCC AGGAACCGTA AAAAGGCCGC GTTGCTGGCG TTTTCCATA GGCTCCGCCC  
 TTGTACTCTC GTTTCCGGT CGTTTTCCGG TCCTTGGCAT TTTTCCGGCG CAACGACCGC AAAAAGGTAT CCGAGGCGGG

4081 CCCTGACGAG CATCACAAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC CAGGCGTTTC  
 GGGACTGCTC GTAGTGT TTTT TAGCTGCGAG TTCAGTCTCC ACCGCTTTGG GCTGTCTGA TATTTCTATG GTCCGAAAG

4161 CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT GCCGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA  
 GGGGACCTTC GAGGGAGCAC GCGAGAGGAC AAGGCTGGGA CGGCGAATGG CCTATGGACA GGCGGAAAGA GGAAGCCCT

4241 AGCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTCGTTTCG TCCAAGCTGG GCTGTGTGCA  
 TCGCACCGCG AAAGAGTTAC GAGTGCAGACA TCCATAGAGT CAAGCCACAT CCAGCAAGCG AGGTTCGACC CGACACACGT

4321 CGAACCCCCC GTTCAGCCCG ACCGCTGCGC CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA CACGACTTAT  
 GCTTGGGGGG CAAGTCGGGC TGGCGACGCG GAATAGGCCA TTGATAGCAG AACTCAGGTT GGGCCATTCT GTGCTGAATA

4401 CGCCACTGGC AGCAGCCACT GGTAACAGGA TTAGCAGAGC GAGGTATGTA GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG  
 GCGGTGACCG TCGTCGGTGA CCATTGTCCT AATCGTCTCG CTCCATACAT CCGCCACGAT GTCTCAAGAA CTTACCACC

4481 CCTAACTACG GCTACACTAG AAGGACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA AAAGAGTTGG  
 GGATTGATGC CGATGTGATC TTCCTGTCTA AAACCATAGA CGCGAGACGA CTCGGTCAA TGGAAGCCTT TTTCTCAACC

4561 TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT GGT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
 ATCGAGAACT AGGCCGTTT TTTGGTGGCG ACCATCGCCA CCAAAAAA ACCTTCGTTT TTTT TTTT TTTT TTTT TTTT

4641 AAGGATCTCA AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG TGAACGAAA ACTCACGTTA AGGGATTTTG  
 TTCCTAGAGT TCTTCTAGGA AACTAGAAAA GATGCCCGAG ACTGCGAGTC ACCTTGCTTT TGAGTGCAAT TCCCTAAAAC

4721 GTCATGAGAT TATCAAAAAG GATCTTCACC TAGATCCTTT TAAATTAATA ATGAAGTTT AAATCAATCT AAAGTATATA  
 CAGTACTCTA ATAGTTTTTC CTAGAAGTGG ATCTAGGAAA ATTTAATTTT TACTTCAAAA TTTAGTTAGA TTTCATATAT

4801 TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA TCTCAGCGAT CTGTCTATTT CGTTCATCCA  
 ACTCATTTGA ACCAGACTGT CAATGGTTAC GAATTAGTCA CTCCGTGGAT AGAGTCGCTA GACAGATAAA GCAAGTAGGT

4881 TAGTTGCCTG ACTCCCCGTC GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC CAGTGCTGTC AATGATACCG  
 ATCAACGGAC TGAGGGGCG CACATCTATT GATGCTATGC CCTCCCGAAT GGTAAGCCGG GTTACGACG TTTACTATGGC

4961 CGAGACCCAC GCTCACCAGC TCCAGATTTA TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCTGTC  
 GCTCTGGGTG CGAGTGGCCG AGGTCTAAAT AGTCGTTATT TGGTCCGGTC GCCTTCCCGG CTCGCGTCTT CACCAGGACG

5041 AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA  
 TTGAAATAGG CGGAGGTAGG TCAGATAATT AACAACGGCC CTTGATCTC ATTCATCAAG CCGTCAATTA TCAAACGCGT

5121 ACGTTGTTGC CATTGCTACA GGCATCGTGG TGTACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA  
 TGCAACAACG GTAACGATGT CCGTAGCACC ACAGTGCAG CAGCAAACCA TACCGAAGTA AGTCGAGGCC AAGGGTTGCT

5201 TCAAGGCGAG TTACATGATC CCCATGTTG TGCAAAAAG CGGTTAGCTC CTTGCTCCT CCGATCGTTG TCAGAAGTAA  
 AGTTCCGCTC AATGTACTAG GGGGTACAAC ACGTTTTTTC GCCAATCGAG GAAGCCAGGA GGCTAGCAAC AGTCTTCATT

5281 GTTGGCCGCA GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC TTAGTGTCT GCCATCCGTA AGATGCTTTT  
 CAACCGGCGT CACAATAGTG AGTACCAATA CCGTCGTGAC GTATTAAGAG AATGACAGTA CCGTAGGCAT TCTACGAAAA

FIG. 4E

5361 CTGTGACTGG TGAGTACTCA ACCAAGTCAT TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC GCGTCAATA  
GACACTGACC ACTCATGAGT TGGTTCAGTA AGACTCTTAT CACATACGCC GCTGGCTCAA CGAGAACGGG CCGCAGTTAT

---

5441 CGGGATAATA CCGCGCCACA TAGCAGAACT TTAAGAGTGC TCATCATTGG AAAACGTTCT TCGGGGCGAA AACTCTCAAG  
GCCCTATTAT GCGCGGGTGT ATCGTCTTGA AATTTTCACG AGTAGTAACC TTTTGCAAGA AGCCCCGCTT TTGAGAGTTC

---

5521 GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA  
CTAGAATGGC GACAACTCTA GGTCAAGCTA CATTGGGTGA GCACGTGGGT TGACTAGAAG TCGTAGAAAA TGAAAGTGGT

---

5601 GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC  
CGCAAAGACC CACTCGTTTT TGTCTTCCG TTTTACGGCG TTTTTCCTCT TATTCCCGCT GTGCCTTTAC AACTTATGAG

---

5681 ATACTCTTCC TTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG AATGTATTTA  
TATGAGAAGG AAAAAGTTAT AATAACTTCG TAAATAGTCC CAATAACAGA GTACTCGCCT ATGTATAAAC TTACATAAAT

---

5761 GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC CTGACGTCTA AGAAACCATT ATTATCATGA  
CTTTTATTT GTTTATCCCC AAGGCGCGTG TAAAGGGGCT TTTCACGGTG GACTGCAGAT TCTTTGGTAA TAATAGTACT

---

5841 CATTAACCTA TAAAAATAGG CGTATCACGA GGCCCTTTTCG TC  
GTAATTGGAT ATTTTATCC GCATAGTGCT CCGGGAAAGC AG

---

FIG. 4F

# Sucrose Gradient Centrifugation of Purified r-sAg Particles

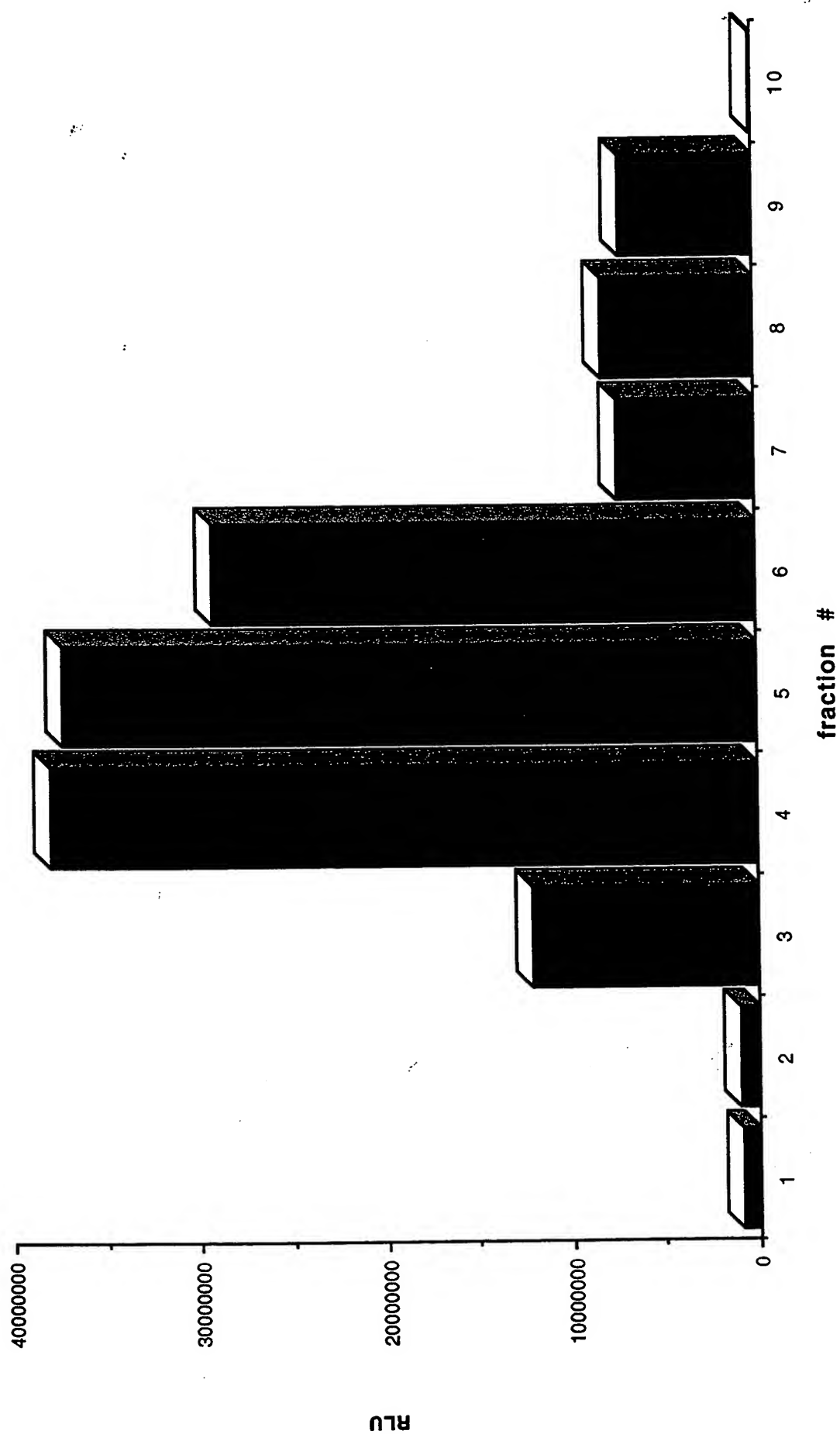


Figure 5

sAg-encoding plasmid titrations  
with E2-sAg Fusion plasmid  
Assay: MAb sAg/MAb E2

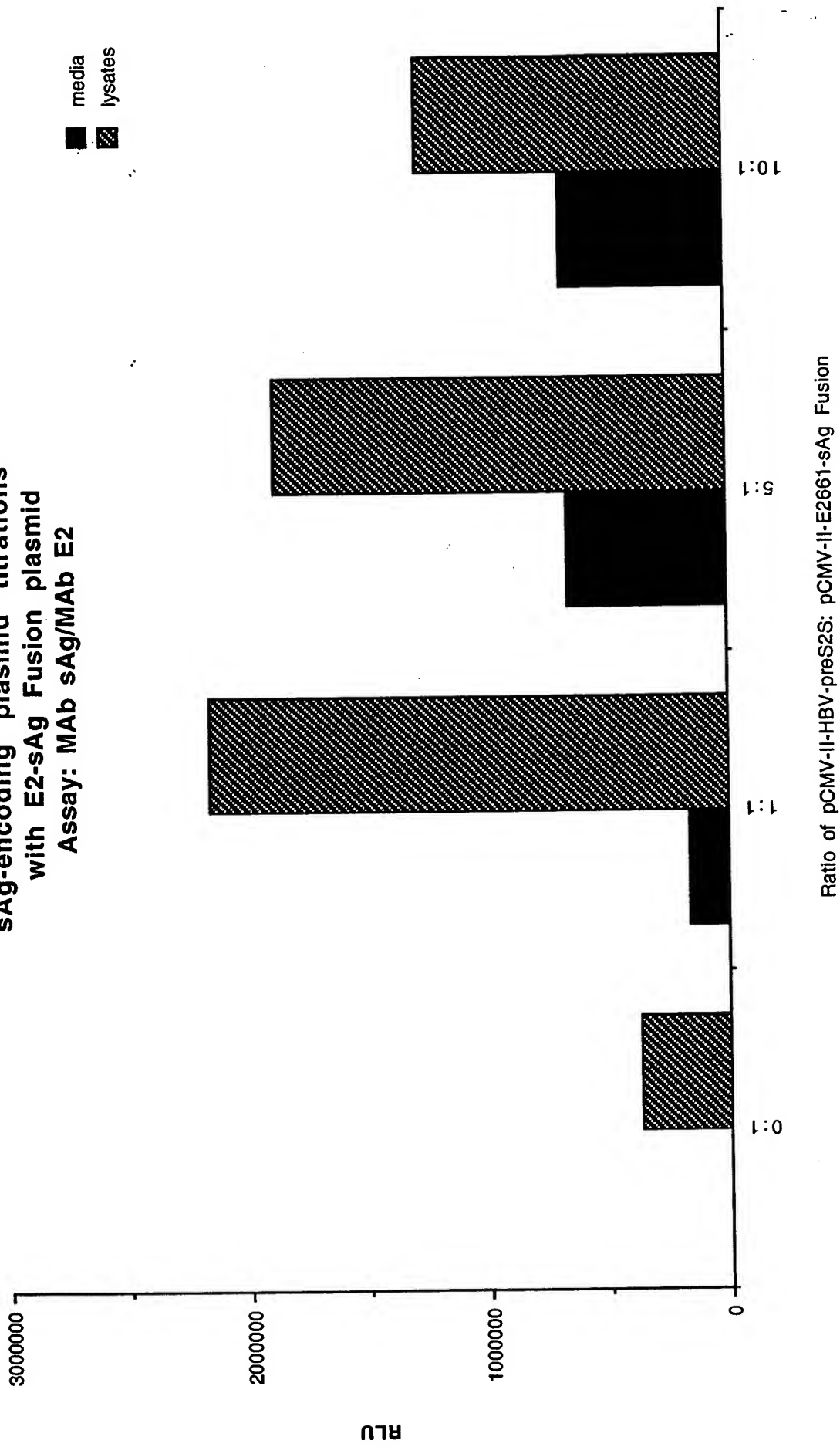


Figure 6A



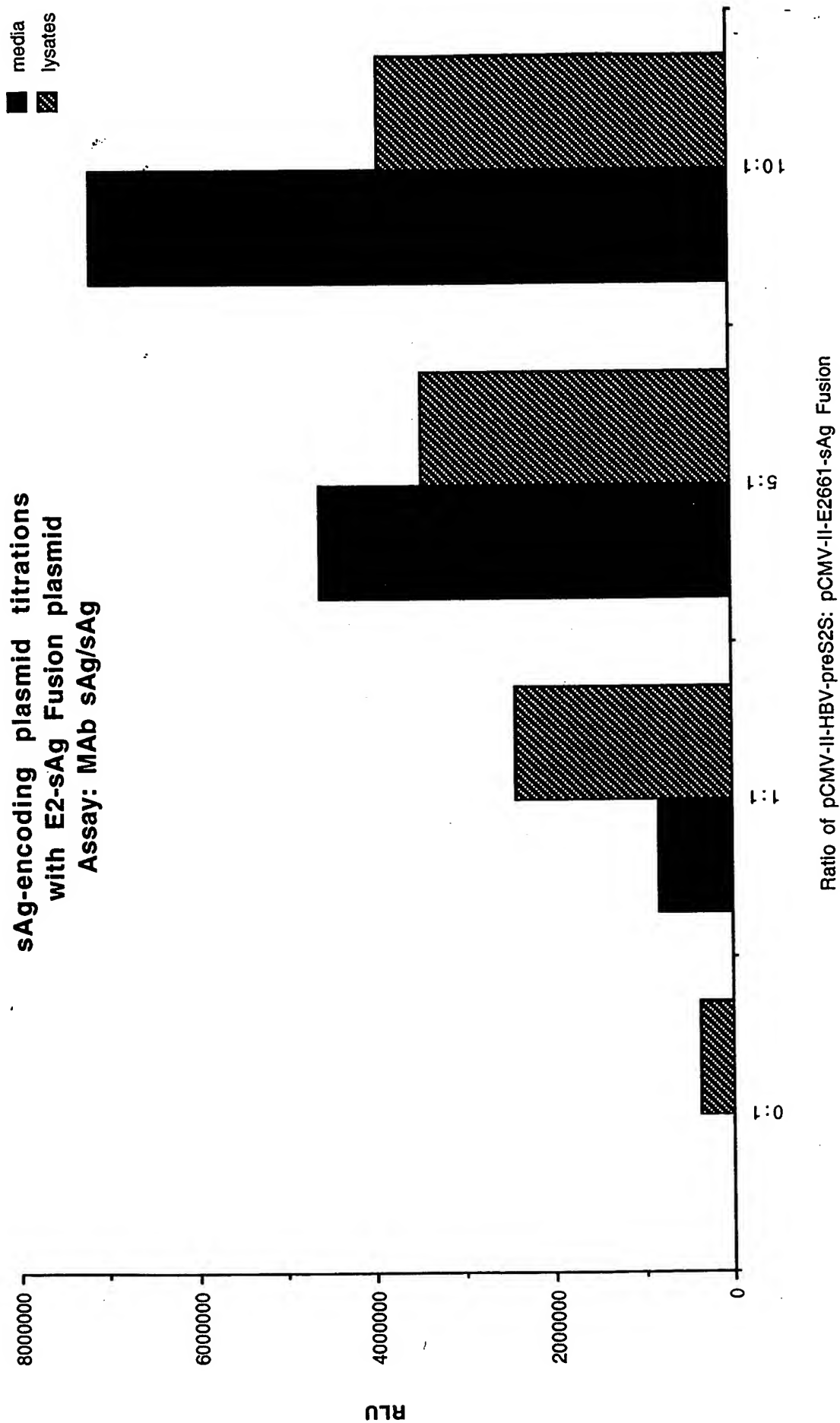


Figure 6B

Sucrose Gradient Sedimentation of Media  
from COS7 Cells Transiently Transfected  
with Different Ratios of sAg:E2 Fusion  
Assay: MAb sAg/ MAb sAg

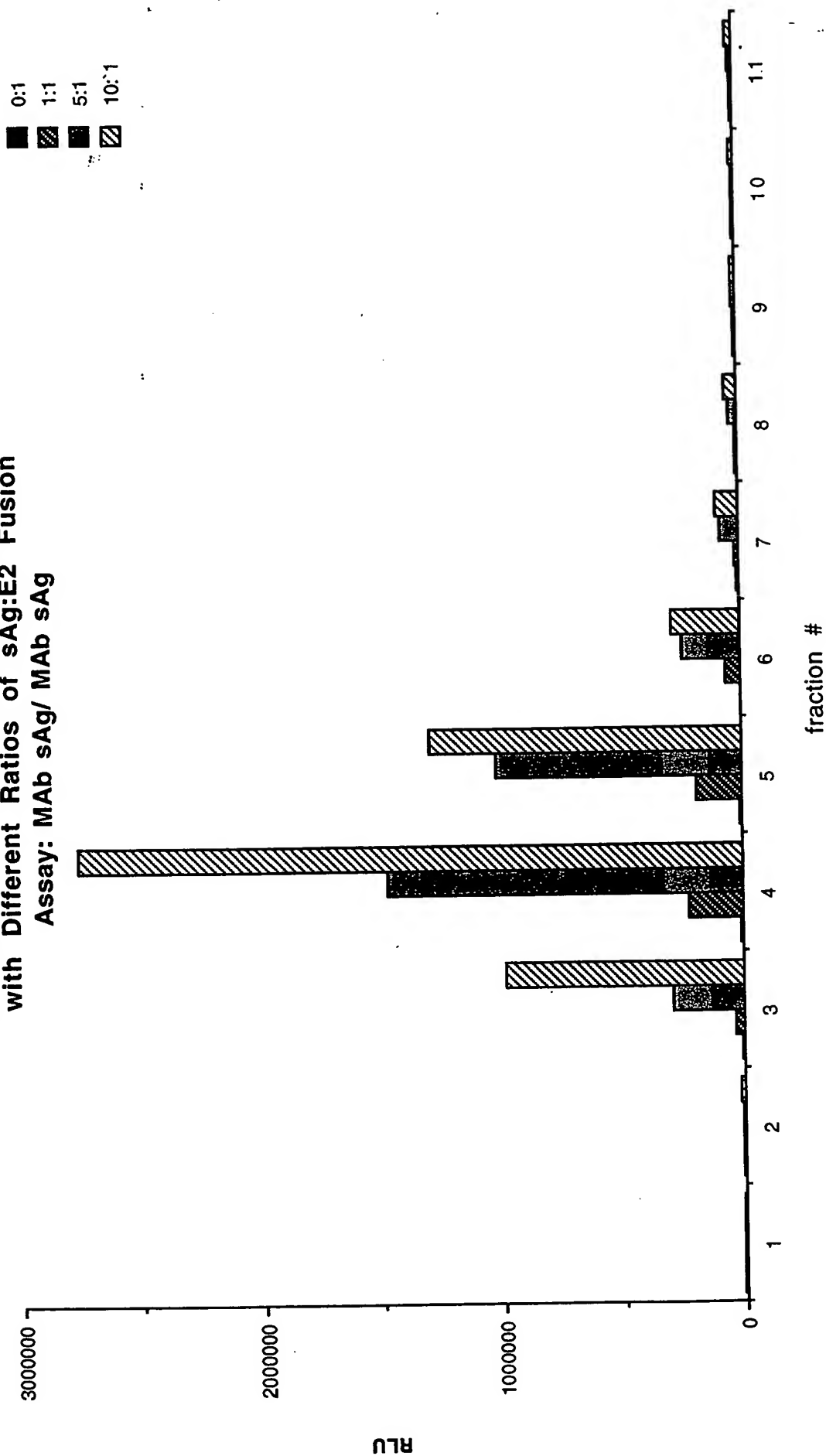


Figure 7A

Sucrose Gradient Sedimentation of Media  
from COS7 Cells Transiently Transfected  
with Different Ratios of sAg:E2 Fusion  
Assay: MAb sAg/ MAb E2

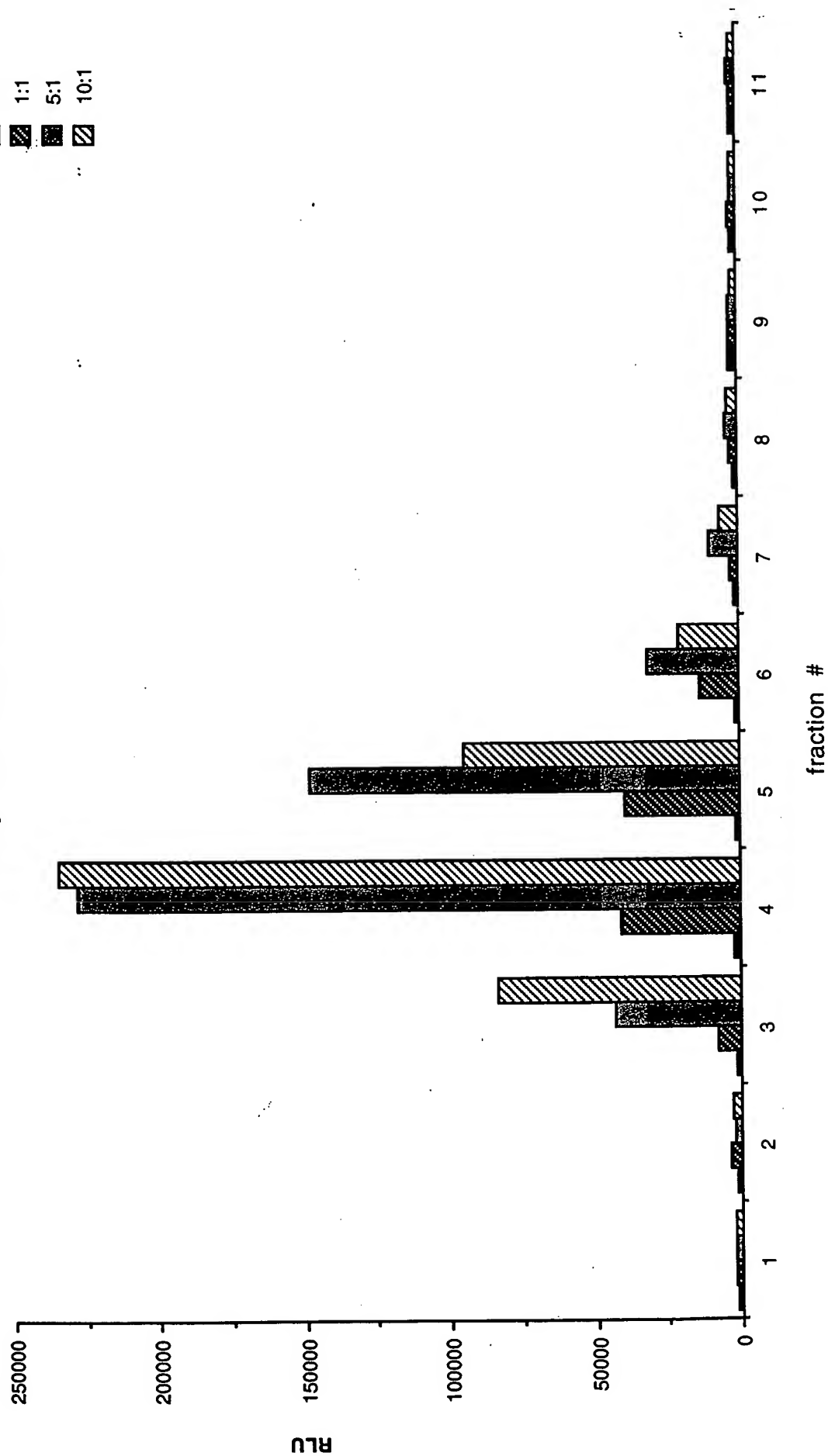
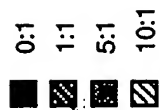


Figure 7B

sAg-ncoding plasmid titrations  
with E1-sAg Fusion plasmid  
Assay: MAb sAg/MAb E1

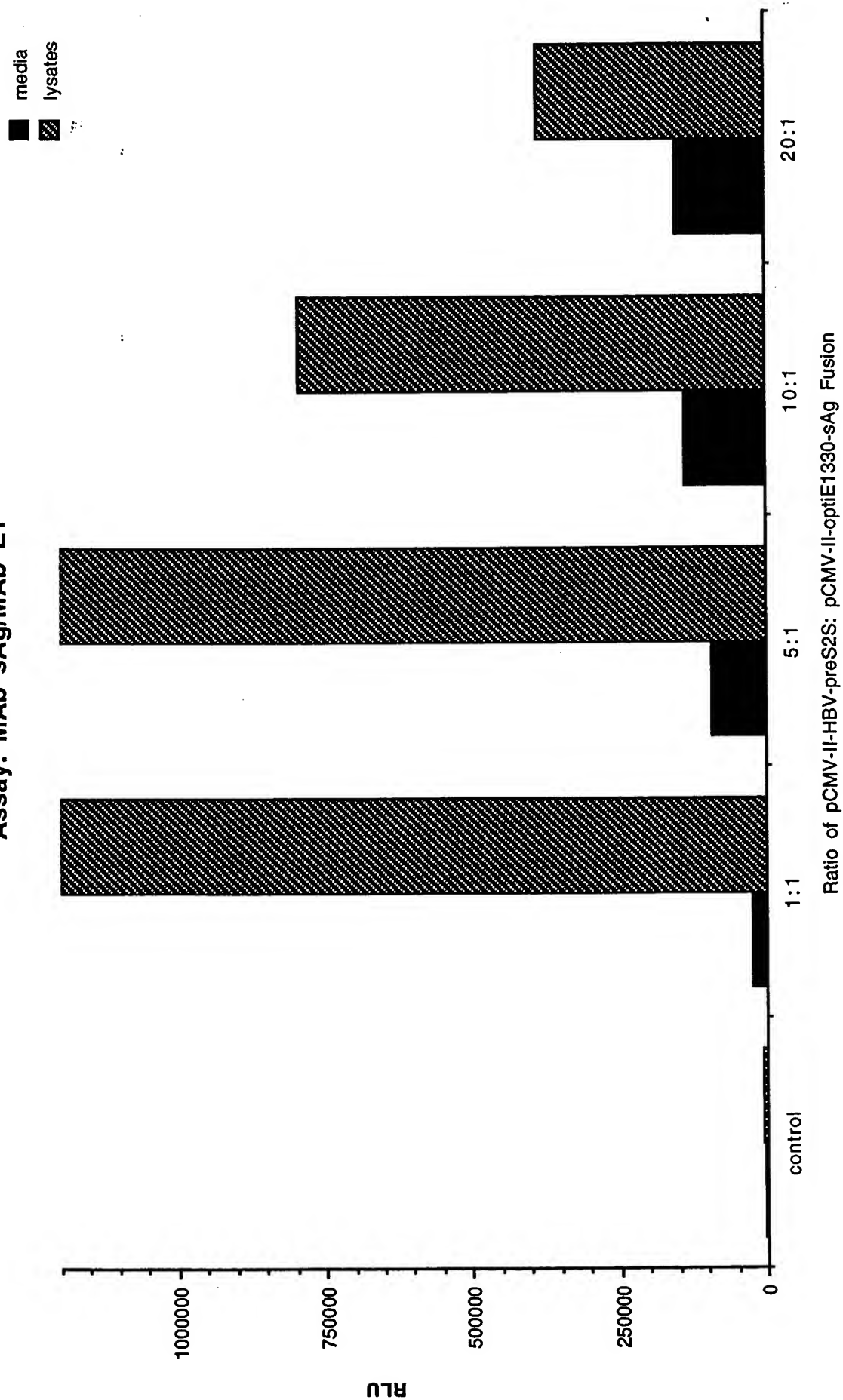


Figure 8A

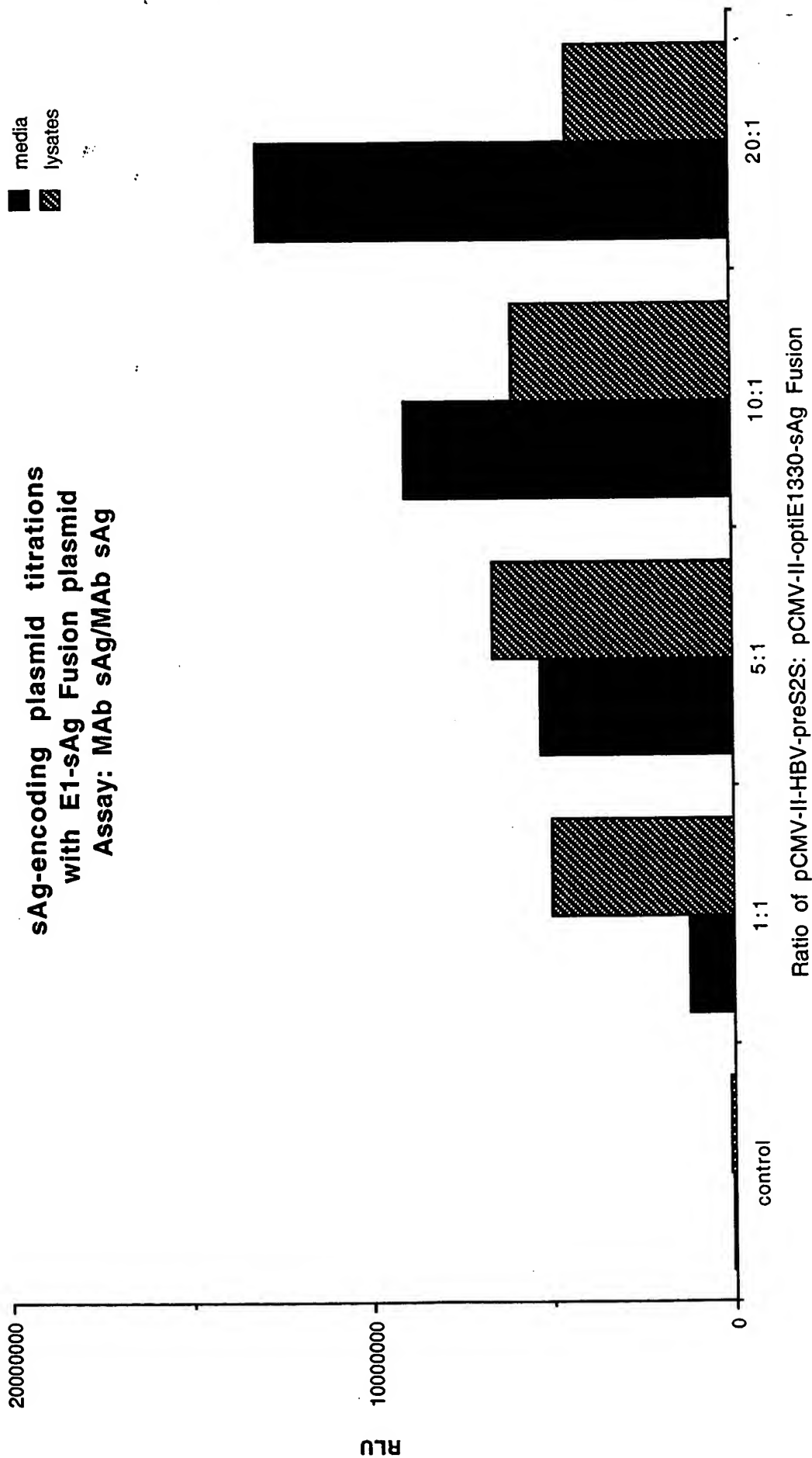


Figure 8B

BEST AVAILABLE COPY

Sucrose Gradient Sedimentation of Media  
from COS7 Cells Transiently Transfected  
with 5x Excess of sAg to E1 Fusion  
Assay: MAb sAg/ MAb E1

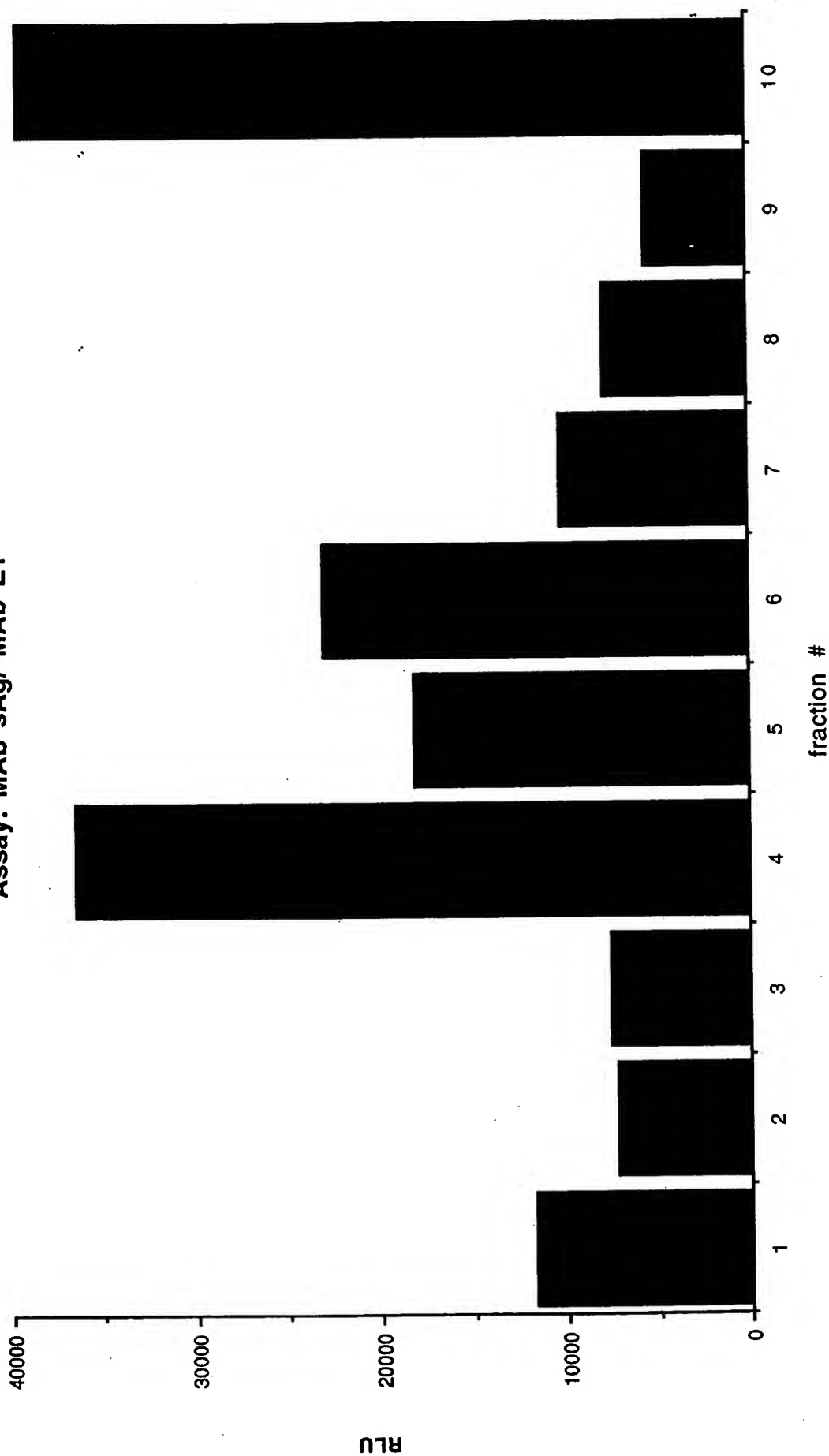
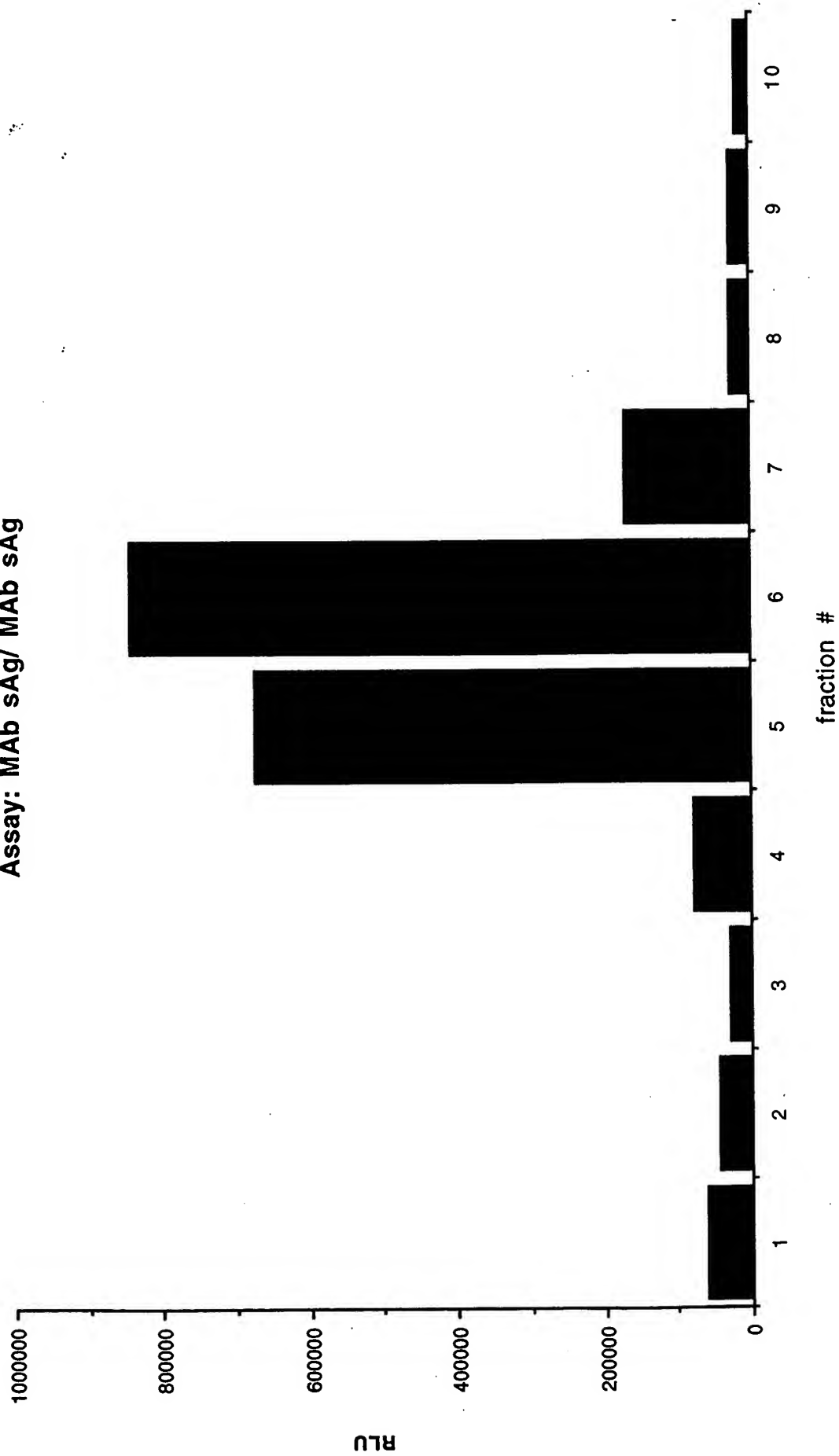


Figure 9A

BEST AVAILABLE COPY

**Sucrose Gradient Sedimentation of Media  
from COS7 Cells Transiently Transfected  
with 5x Excess of sAg to E1 Fusion  
Assay: MAb sAg/ MAb sAg**

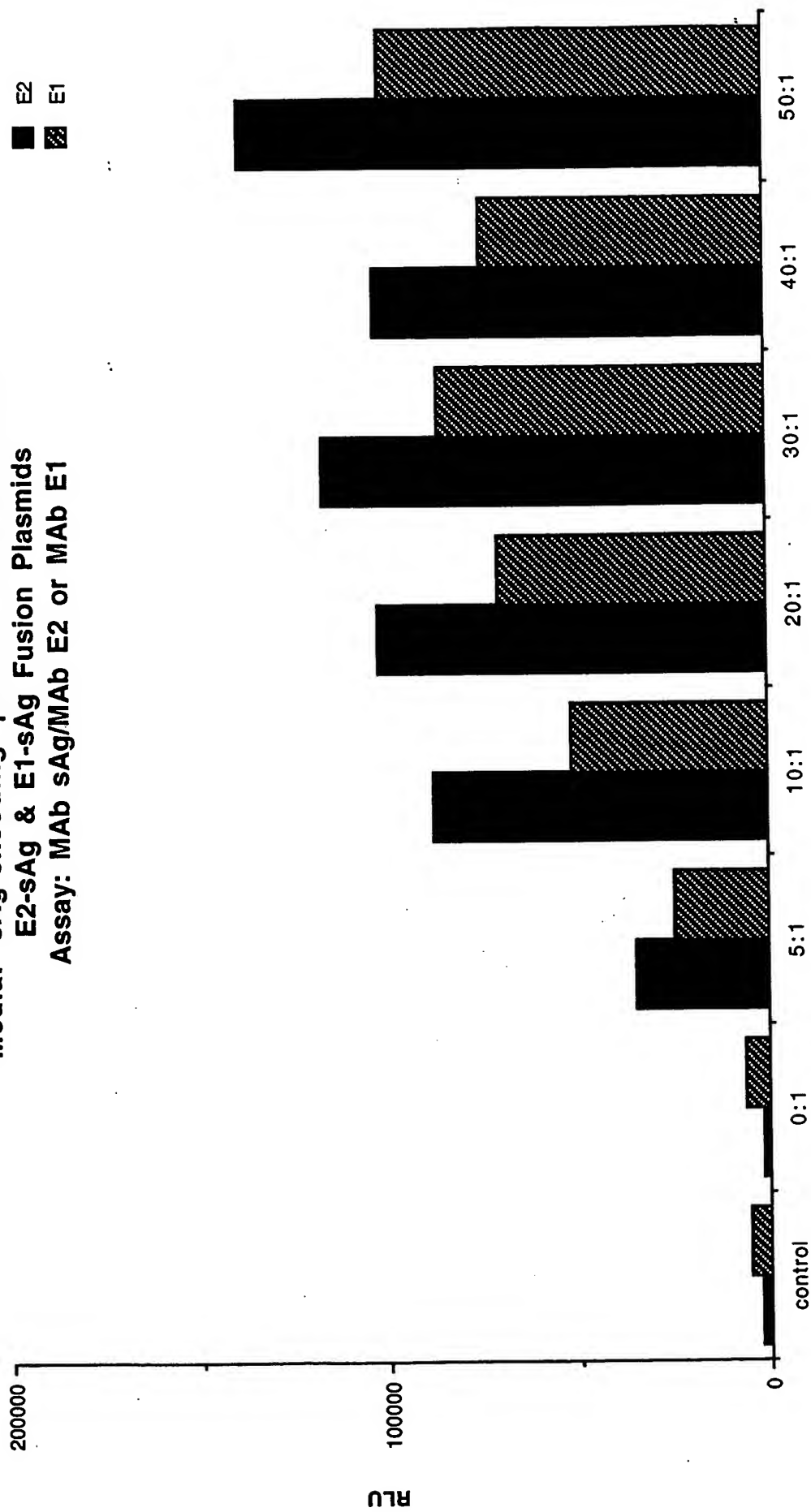


**Figure 9B**

**BEST AVAILABLE COPY**

Media: sAg-encoding plasmid titrations with  
E2-sAg & E1-sAg Fusion Plasmids  
Assay: MAb sAg/MAb E2 or MAb E1

■ E2  
▨ E1



Ratio of pCMV-II-HBV-preS2S to  
pCMV-II-E2661-sAg and pCMV-II-optIE1330-sAg Fusions

Figure 10A



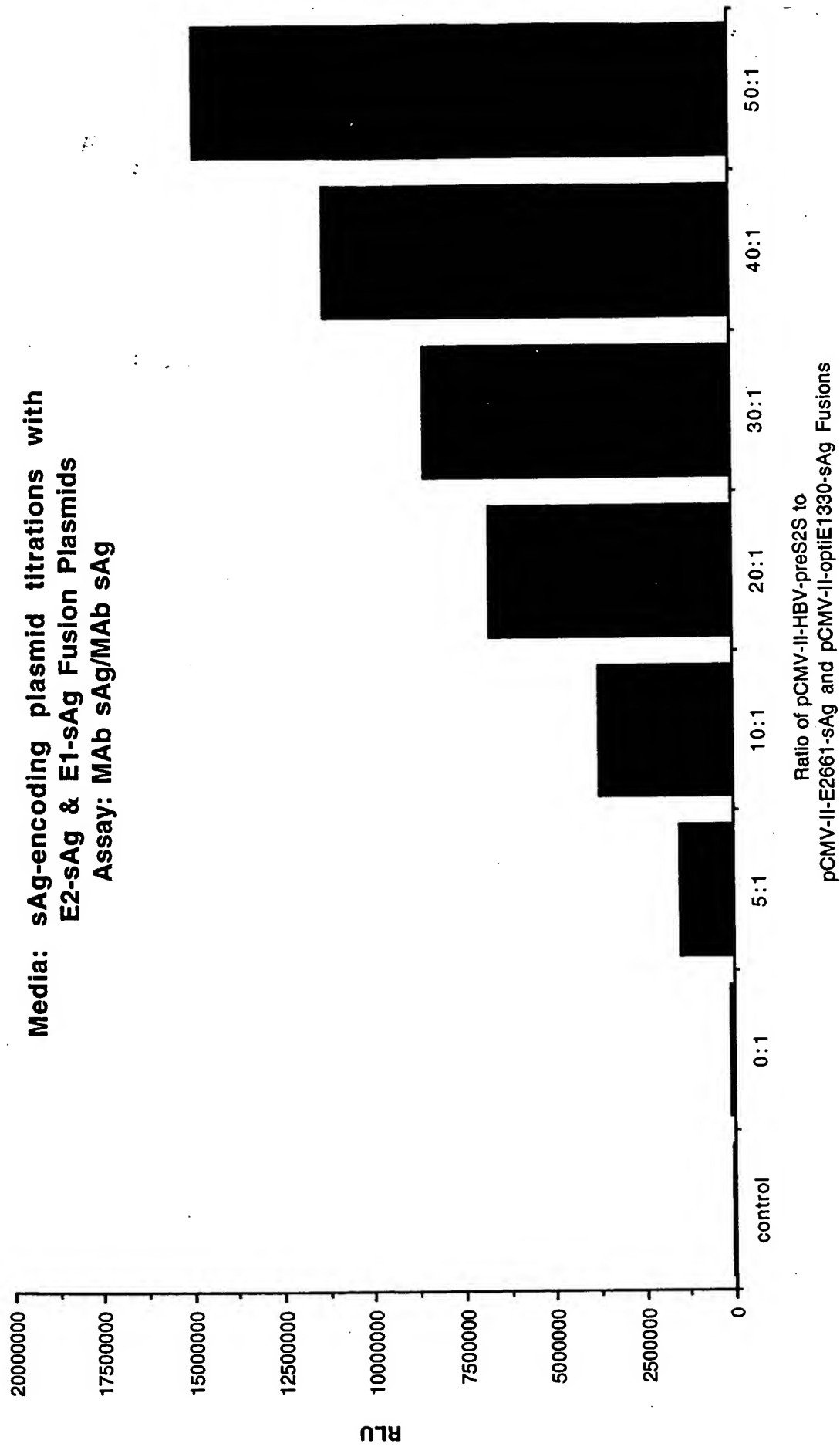


Figure 10B

BEST AVAILABLE COPY